邮 电 高 等 学 校 教 材

概率论与随机过程

许升汉 编

人民邮电出版社

邮电高等学校教材

概率论与随机过程

许升汉 编

人民邮电出版社

内 容 提 要

本书是电信专业的研究生进一步学习概率论和随机过程的一本教材和参考书。全书共分十章。前五章为概率论部分，包括概率空间、随机变量及其分布函数、随机变量的数字特征、随机变量的特征函数、大数定理及中心极限定理。后五章为随机过程部分，内容包括随机过程的基本概念、随机分析、平稳过程、Markov链及可数齐次 Markov 过程等。

邮电高等学校教材

概率论与随机过程

许升汉 编

人民邮电出版社出版

北京朝阳门内南竹杆胡同 111 号

中国铁道出版社印刷厂印刷

新华书店北京发行所发行

各 地 新 华 书 店 经 售

*

开本：850×1168 /32　　1996 年 5 月　　第一版

印张：14　　1996 年 5 月　　第 1 次印刷

字数：371 千字　　印数：1 — 1000 册

ISBN 7-115-05825-3/F・136

定价：13.50 元

编 者 的 话

本书是为了工科院校的研究生及工科院校毕业生进一步学习概率论及随机过程而编写的。

考虑到读者自学的需要，所讲述的问题都尽可能给出详细的推导。

本书是编者在 1978 至 1991 年期间在北京邮电大学及中国科学院科技大学研究生院讲授本课程的讲稿的基础上修改编写而成的。由于学时的限制，本书中仅给出了在信息论、信号检测、图像处理等课程所用到的必不可少的内容，并为进一步学习这方面的知识打下必要的基础。由于版幅限制，许多证明只能从略，请谅解。

由于编者水平有限，不论在内容选材上及问题的论述上都不免有不足之处，恳请读者指正。

在本书编写过程中，曾得到蔡宗蔚教授、姜炳麟教授、王玉孝副教授多方指导和帮助，特此致谢。

编　者

1995 年

目　录

第一章 概率空间

本章的主要目的是在已学过的工科院校概率论教材的基础上，建立客观事物中随机事件及其概率的数学模型——概率空间。从而使读者对随机现象的本质有进一步的理解，并为进一步的学习及研究随机现象打下基础。

为叙述简便，用 E 表示随机试验。即在一组可以重复实现的条件下，若所观察的现象在每一次观察中可能出现也可能不出现。这种观察过程，称为随机试验，以后简称为试验。

试验 E 中每一个可能的结果称为样本点，用 ω 或带有下标的 ω_i 等表示之。试验 E 中一切可能结果的集合称为样本空间，记成 Ω_E。在没有误会的条件下，Ω_E 又简记为 Ω，集合 Ω 可表为

$$\Omega=\{\omega:\omega \text{ 是试验 } E \text{ 中的可能结果}\}$$

在研究某随机问题时，往往要通过试验 E 去观察某种现象（或随机事件），而所观察的现象往往是由 Ω 中的若干样本点所构成的集合，由此可见，试验 E 中的随机事件（以后简称为事件）一定是该试验 E 所确定的样本空间 Ω_E 的子集。若 A 表示试验 E 中的事件，按集合的表示方法，则有：

$$A\subset\Omega$$

在对试验 E 的一次观察中，若发现其样本点 ω 出现，且 ω 是 A 中的元素，即发现：

$$\omega\in A$$

则称事件 A 在这次试验中是发生了。

在集合论中，一个集合 Ω 可以认为是它本身的子集，即：

$$\Omega\subset\Omega$$

这样的子集 Ω 就包含了试验 E 中一切可能的结果，因此每一次试验

Ω 必发生，因此称 Ω 是试验 E 中的必然事件。

$$A^C \triangleq \Omega - A$$

称为 A 事件的余事件(或称为对立事件)。

$$\omega \in A, \qquad 则 \ \omega \overline{\in} A^C$$

因此在一次试验中，A、A^C 中有一且必有一发生。这样就有：

$$\Phi \triangleq \Omega^C = \Omega - \Omega$$

其中 $\triangleq$ 表示等式左边由等式的右边去定义。

显然，Φ 不含任何试验 E 中的结果，亦即 Φ 中不含 Ω 中的任何样本点。在集合论中，若所讨论的范围是集合 Ω_E 中的元素，Φ 不含 Ω_E 中的任何元素，则称 Φ 是所研究的范围内的空集，简称为空集。由于 Φ 中不含 Ω_E 中任何样本点，因此任何一次试验 E，事件 Φ 均不发生，因此在概率论中称 Φ 为不可能事件。

以上介绍了本书中常用到的一些术语及符号。下面正式讲述本章的内容。

§1.1 概率空间的概念

在目前工科院校的概率课程中，对随机事件仅给出描述性的说明，对事件的概率仅在古典概型中给出了明确的定义，但在实际的各种随机现象中，显然并不能全用古典概型去概括。例如当 Ω 中含有可列多个样本点时，就无法用古典概型计算事件的概率的公式去计算这种试验中事件的概率。为了找出各种复杂随机现象中随机事件及其概率的一般的、准确的描述方法，就需要从对各种具体的随机现象的研究中，找出它们的共性，从而建立起一个严格的并能概括所有随机现象及其概率的数学模型，使对随机现象的研究建立在科学的基础之上。为此，先研究三类具体的随机现象的模型及其主要性质，以便从中概括出具有普遍意义的随机现象的数学模型。

一、三种典型的随机问题

1. 古典概型

古典概型是在古代较早的时候，在某些特殊情况下，利用研究对象(如掷硬币、掷骰子)在物理和几何性质所具有的均匀性、对称性去计算所观察的随机现象出现可能性大小的一种方法，它后来被较精确地概括为：

<u>定义 1.1.1</u>　试验 E 若满足：

(1)　其一切可能的结果是有限多个。其一切可能的结果用 ω_1, $\omega_2, \cdots, \omega_n$ 表示，即

$$\Omega_E = \{\omega_1, \omega_2, \cdots, \omega_n\}$$

(2)　用 $\{\omega_k\}, k=1,2,\cdots,n$ 表示只含有一个可能结果(或样本点)的事件，称 $\{\omega_k\}$ 是基本事件。用 $P(\{\omega_k\})(k=1,2,\cdots,n)$ 表示基本事件在试验 E 中出现的可能性的大小或是基本事件的概率，则有

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \cdots = P(\{\omega_n\}) = \frac{1}{n}$$

这个特性称为等可能性，它是对研究对象在质量的均匀性和物体几何形状的对称性的概括。

(3)　若所观察的随机事件用 A 表示，且 A 中含有 m_A 个样本点，则规定事件 A 在试验中出现的概率 $P(A)$ 为：

$$P(A) \triangleq \frac{m_A}{n} = \frac{A\text{中所含样本点数}}{\Omega\text{中所含样本点数}} \tag{1.1.1}$$

则称试验 E 是一古典概型的试验，或简称为古典概型。式(1.1.1)就是读者所熟悉的古典概型事件概率的计算公式。

古典概型中的事件及其概率有下述重要性质：

<u>定理 1.1.1</u>　设 E 是一古典概型的试验，则有：

(1)设 A 是试验 E 中的任一事件，则恒有：

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

即事件发生可能性的大小是一非负的且不超过 1 的实数。

(2) $$P(\Omega)=1$$
即 E 中的必然事件的概率是 1。

(3)若 $A_1,A_2,\cdots,A_m$ 是试验 E 中两两互不相交的事件，即$A_i\cap A_j=\Phi(i\neq j,i,j=1,2,\cdots,m)$。用 $\sum\limits_{k=1}^{m}A_k$ 表上述 m 个两两互不相交事件的和，若 $\sum\limits_{k=1}^{m}A_k$ 是试验 E 中的事件，则有：

$$P(\sum_{k=1}^{m}A_k)=\sum_{k=1}^{m}P(A_k)$$

亦即：当有限个两两不相交的试验 E 中事件之和也是试验 E 中的事件时，则其和事件发生的概率概等于每一个事件发生的概率之和。(此定理读者可自行证明)。

2. 几何概型

为了直观起见，先研究下列具体问题。

例 1.1.1 (约会问题)

有某二人，约定于某日 0 到 T 时内在某地会面，并约定在约定时间内先到达者等待 $t(\leqslant T)$时后，若不见对方来就离去。试研究二人能会面的事件 A 的概率。

由题意知，二人能会面的事件是具有随机性的。如何计算 $P(A)$?

可以设想另有一观察者在二人约会地点记录二人到达约会地点的时间。若用 x、y 分别表示二约会人到达约会地点的时间，数偶(x,y)就是观察者所观察到的一个样本点。显然 x、y 分别满足：$0\leqslant x\leqslant T$，$0\leqslant y\leqslant T$。若我们用一平面直角坐标系(如图 1.1.1)上的点表示这些样本点，则此试验的样本空间 Ω 就是图 1.1.1 中边长为 T 的正方形中全部点的集合，即：

$$\Omega=\{(x,y):0\leqslant x\leqslant T,0\leqslant y\leqslant T\}$$

(在集合论中，填满一直线段和填满如上述正方形中之点的总数称为不可数无穷多个)。

由于 x、y 分别表示二人到达约会地点的时间，若二人能会面，则由约定的条件可知，只有当 x、y 满足下列条件时：

$$|x-y|\leqslant t,(0\leqslant x\leqslant T,0\leqslant y\leqslant T)$$

二人会面的事件 A 才能发生，这样 A 中所含样本点为：

$$A=\{(x,y):|x-y|\leqslant t,0\leqslant x\leqslant T,0\leqslant y\leqslant T\}$$

即图 1.1.2 中阴影部分。

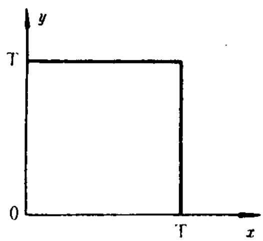

图 1.1.1

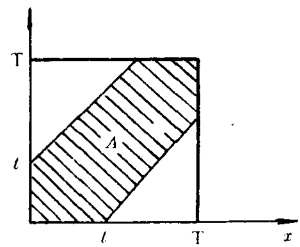

图 1.1.2

显然，Ω 及 A 中都含有不可数多个样本点，因此古典概型中计算事件概率的公式(1.1.1)不再适用于这个问题。

但由直观上可以看出：当 A 在 Ω 中所占的面积愈大，则二人会面的可能性也就愈大。若用 $L(\Omega)$、$L(A)$分别表示样本空间 Ω 及二人能会面的事件 A 在平面上所占的面积，则事件 A 发生的概率为：

$$P(A)\triangleq\frac{L(A)}{L(\Omega)}=\frac{T^2-(T-t)^2}{T^2}=1-(1-\frac{t}{T})^2 \qquad (1.1.2)$$

显然当 t 值越接近于 T 时，二人会面的概率也就越大，而当 t 值越小时，则二人会面的概率也就越小，这完全符合情理，可见式(1.1.2)的规定是合理的。

为了概括出这类问题的特点，下面对例 1.1.1 的问题作进一步分析：

(1) 它的样本空间中全部样本点能构成一个有长度或有面积或有体积的几何图形(如直线段、曲线段、平面图形、空间曲面、空间体)，为方便起见，今后用 $L(\Omega)$表示 Ω 图形的长度、面积或体积，且

$L(\Omega)>0$。

(2) 所考察的事件 A 中全部样本点也构成相应的有长度、面积或体积的图形，并用 $L(A)$表示 A 图形的长度、面积或体积。

(3) 规定比值：

$$P(A) \triangleq \frac{L(A)}{L(\Omega)} \tag{1.1.3}$$

且其 $P(A)$仅与 $L(A)$的大小有关，而与 A 的形状及 A 在 Ω 的图形中所处位置无关。

满足上述三条的随机问题称为几何概型问题，简称为几何概型。

请注意，在这一段里没有称上述三条所描述的随机问题为几何概型的定义，这是因为几何概型的严格定义要用到集合的测度知识，由于我们假定读者不具有集合测度的知识，因而只用了不具备集合测度知识的读者易于接受的语言来描述几何概型。但这并不妨碍对后述概率空间问题的理解。

下面进一步介绍几何概型的性质。

定理 1.1.2　设 E 是一几何概型的试验，Ω 表示试验 E 中的样本空间，且 $L(\Omega)>0$，A 及 $A_k(k=1,2,3,\cdots)$都是试验 E 中的事件，且 $L(A)>0$ 及 $L(A_k)>0(k=1,2,\cdots)$均存在，则有：

(1) 对任何试验 E 中的事件 $A(\subset\Omega)$，均有：

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

(2) $P(\Omega)=1$；

(3) 对可列无穷多个(即能与自然数集$\{1,2,3,\cdots\}$中元素建立起一一对应关系的)事件列 A_1、A_2、A_3、$\cdots$，若

$$A_i\cap A_j=\Phi \quad (i\neq j,i,j=1,2,3,\cdots)$$

且 $\sum_{i=1}^{\infty} A_i$ 仍是 E 中的一事件时，则有：

$$P(\sum_{k=1}^{\infty} A_k)=\sum_{k=1}^{\infty} P(A_k)$$

成立。

本定理中的结论(1)、(2)是显然的，结论(3)称为事件概率的可

列可加性，它的论证要用点集的勒贝格(Lebesgue)测度，此处证明从略。

3. 试验统计方法

前述的两种随机概型显然远远概括不了客观上的复杂的随机现象。当遇到前述两种随机概型所概括不了的随机问题时，在实际工作中往往就采用试验统计的方法，即用事件的频率近似地去计算和表达事件的概率。因此，事件的概率及其性质可以通过事件的频率来理解，为此，先介绍事件频率的定义及性质。

定义 1.1.2　设 E 是一试验，Ω 是其样本空间，$A(\subset\Omega)$是试验 E 中的任一事件。

若在同样的条件下，将试验 E 重复做 n 次，事件 A 在 n 次试验中出现了 μ_A 次，则称 μ_A 是这 n 次试验中事件 A 发生的频数.或简称为频数，比值：

$$f_A^{(n)} \triangleq \frac{\mu_A}{n}$$

则称 $f_A^{(n)}$ 为 n 次试验中事件 A 发生的频率或简称为事件 A 的频率。

显然，事件 A 的频率是具有随机性的。同时试验表明事件的频率在客观上还具有稳定性，即在试验次数 n 增大时，其中大量的频率数值聚集在一个常数的周围。此常数是客观存在的，它反映了事件 A 出现的可能性的大小，这个常数称为事件 A 的概率。仍用符号 $P(A)$ 表示。

由此可见事件的频率在一定程度上反映了事件概率的特征，事件频率的性质也在客观上反映了事件概率应具有的性质。

定理 1.1.3　对任意的试验 E，事件的频率具有下列性质：

(1) 对任一事件 A，均有：

$$0\leqslant f_A^{(n)}\leqslant 1$$

(2) $f_\Omega^{(n)}=1$；

(3) 设 $A_k(k=1,2,\cdots,m)$是试验 E 中两两互不相交的 m 个事件，

若 $\sum_{k=1}^{m} A_k$ 仍是 E 中的事件，则有：

$$f^{(n)}_{\sum_{k=1}^{m} A_k} = \sum_{k=1}^{m} f^{(n)}_{A_k}$$

此定理读者可自行证明。

综合上述关于各类随机问题的研究就会发现：

（1）不同类型的随机问题中事件的概率的计算方法是不同的，且无法统一其计算公式。因此要想得到随机现象中事件及其概率的统一模式，就不能由统一其计算方法入手。

（2）虽然不同类型的随机问题中事件的概率的计算公式可以是各式各样的，但进一步分析就会发现它们有着下列共同的特点：

①事件是样本空间的子集；

②事件的概率 $P(A)$是事件 A 的函数。由于事件 A 是 Ω 的子集，因此，$P(A)$是以 Ω 中部分子集为变元的集合函数；

③表达事件概率的集合函数均满足：

（a）非负性，即对任一事件均有：

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

（b）正规性，

$$P(\Omega)=1$$

（c）可列可加性，即两两不相交的可列多个事件 $A_1, A_2, \cdots, A_k \cdots$，当 $\sum_{k=1}^{\infty} A_k$ 仍是一事件时，则必有：

$$P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$$

由上述分析可知，事件的概率 $P(A)$可用定义域为 Ω 中部分子集，其函数值域为[0,1]且满足非负性、正规性及可列可加性的集合函数来定义，但在给出一般性的定义之前，还必须弄清楚其定义域的结构。

二、集合的 σ—代数

作为事件的概率的这种集合函数的定义域有什么特点呢？

首先可设想，将试验中的所有的事件取出来，以这些事件为元素构成一新的集合，在集合论中称以集合为元素的集合为集合类，以后用 $\mathscr{F}$ 表示试验 E 中一切事件构成的集合类，即：

$$\mathscr{F} \triangleq \{A: A\text{ 是 }E\text{ 中的事件}\}$$

显然 $\mathscr{F}$ 就是试验 E 中表达事件概率的集合函数 $P(A)$ 的定义域。那么 $\mathscr{F}$ 有什么特点呢？

(1) 由于 $\Omega\subset\Omega$，且 Ω 表示 E 中的必然事件，从而必有：

$$\Omega\in\mathscr{F}$$

(2) 当 $A\in\mathscr{F}$ 时，由试验表明，每一次试验的结果 ω 必有：

$$\omega\in A \qquad \text{或 } \omega\in A^c$$

这样就有：

若 $A\in\mathscr{F}$，则必有：

$$A^c\in\mathscr{F}$$

(3) 再由于 $P(A)$ 应具有(有限)可列可加性，这就要求，当

$$A_k\in\mathscr{F}(k=1,2,\cdots) \quad A_i\cap A_j=\Phi \quad (i\neq j)$$

时，则应有：

$$\sum_{k=1}^{n}A_k\in\mathscr{F} \quad \text{及} \sum_{k=1}^{\infty}A_k\in\mathscr{F}$$

否则 $P(A)$ 的可列可加性将变得毫无意义了。

这样我们将 $\mathscr{F}$ 所应具有的上述三条性质稍加推广，就得到下述定义：

定义 1.1.3　设 Ω 是任一集合，$\mathscr{F}$ 是由 Ω 中部分子集所构成的集合类，若 $\mathscr{F}$ 满足：

(1) $\Omega\in\mathscr{F}$；

(2) 若 $A\in\mathscr{F}$，则 $A^c\in\mathscr{F}$；

(3) 若 $A_k\in\mathscr{F}(k=1,2,\cdots,n)$

则
$$\bigcup_{k=1}^{n} A_k \in \mathscr{F}$$
称 $\mathscr{F}$ 是 Ω 上的一个集合代数;

(4) 若 $A_k \in \mathscr{F} \quad (k=1,2,\cdots)$

则
$$\bigcup_{k=1}^{\infty} A_k \in \mathscr{F}$$
称 $\mathscr{F}$ 是 Ω 上的一个 σ—代数。

显然,定义 1.1.3 中(3)、(4)是前述分析中(3)的推广。当定义 1.1.3 中(3)、(4)满足时前述分析中(3)必满足。

代数与 σ—代数不同之处在于:集合代数仅对有限并运算封闭,而 σ—代数则对可列并运算封闭。

例 1.1.2 设 Ω 是任一集合,若
$$\mathscr{F} \triangleq \{A, A \subset \Omega\}$$
即 $\mathscr{F}$ 是由 Ω 中一切子集所构成的集合类,则 $\mathscr{F}$ 是一 σ—代数。

例 1.1.3 设 Ω 是任一集合,$A \subset \Omega$,若
$$\mathscr{F} \triangleq \{\Omega, A, A^C, \Phi\}$$
则 $\mathscr{F}$ 是一 σ—代数。

定理 1.1.4 设 $\mathscr{F}$ 是 Ω 上的 σ—代数,则 $\mathscr{F}$ 具有下列性质:

(1) $\Phi \in \mathscr{F}$;

(2) 若 $A_k \in \mathscr{F} \quad (k=1,2,\cdots)$,则 $\bigcap\limits_{k=1}^{\infty} A_k \in \mathscr{F}$,即 σ—代数对可列交运算也封闭;

(3) 设 $A_k \in \mathscr{F} (k=1,2\cdots,n)$,则 $\bigcup\limits_{k=1}^{n} \in \mathscr{F}$,即 σ—代数对有限并运算封闭,从而 σ—代数必是代数。

(4) 设 A、$B \in \mathscr{F}$,则
$$A - B \in \mathscr{F}$$
即 σ—代数对集合的差运算是封闭的。

证明:由 σ—代数的定义中条件(1)可知 $\Omega \in \mathscr{F}$,再由定义中条件(2),则有:
$$\Phi = \Omega^C \in \mathscr{F}$$