计算数学讲义（六）

概率统计基础和概率统计方法

南京大学数学系计算数学专业 编

科学出版社

计算数学讲义(六)

概率统计基础和概率统计方法

南京大学数学系计算数学专业 编

科学出版社

1979

内 容 简 介

本书为综合性大学理科计算数学专业通用教材，全书分为两部分，共七章．第一部分包括第一、二两章，分别介绍概率论和数理统计的基础知识，并附有习题．第二部分包括第三至第七章，介绍数理统计的一些方法：第三、四章分别介绍方差分析和回归分析方法，第五章为逐步回归方法，第六章介绍判别分析、逐步判别以及 K-均值聚类分析等方法，第七章对统计试验法作了简单的介绍．

本书可供计算数学工作者及有关工程技术人员参考．

215/75

计算数学讲义(六)

概率统计基础和概率统计方法

南京大学数学系计算数学专业 编

*

科学出版社 出版

北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1979年 8月第 一 版　　开本：787×1092 1/32

1979年 8月第一次印刷　印张：11 7/8　插页：1

印数：0001—83,160　　字数：272,000

统一书号：13031・920

本社书号：1304・13—1

定价： 1.25 元

说　　明

一、这一套《计算数学讲义》是在我专业过去所编教材的基础上修改补充而成的.

二、这套讲义共分下列九册:

（一）数值逼近方法;

（二）线性代数计算方法;

（三）常微分方程数值解法;

（四）偏微分方程数值解法;

（五）最优化方法;

（六）概率统计基础和概率统计方法;

数学基础之一: 线性代数;

数学基础之二: 常微分方程;

数学基础之三: 偏微分方程.

三、这套讲义可作为综合性大学理科计算数学专业教材,也可供利用电子计算机从事科学计算的科技人员参考.

四、这套《计算数学讲义》的主编是何旭初同志.

讲义各册由我专业有关同志分工负责.

这册《概率统计基础和概率统计方法》的编写者为唐述钊、屠俊如和王嘉松同志.

五、由于理论水平和实践经验所限，讲义中的缺点和错误在所难免,我们衷心盼望读者提出宝贵意见,以便进一步修改.

南京大学数学系计算数学专业

1978年1月

目　　录

第一部分　概率论和数理统计基础

第二部分　数理统计方法

第一部分　概率论和数理统计基础

第一章　概率论基础

§1. 概率空间

1.1. 引言　在现实世界中，许多现象我们是无法事先断定它们的结果的．例如，我们知道气体是由无数气体分子组成的，每个分子不断地运动着，但对于分子在一特定时刻的位置、速度、方向等我们却不能精确地断定．又如在任一分钟内，一个电话交换台是否接到呼唤以及有多少次呼唤等也是无法事先断定的．这类现象(或试验)我们称之为随机的．尽管个别的随机现象是无规律的，但大量的随机现象(或在同一条件下大量的重复试验)却显现着某种规律性．正如个别气体分子运动是随机的，但一容器内气体对器壁的总影响(压力)却是稳定的．也就是说大量气体分子运动时所显现出的总现象、压强等是有规律的．而概率论就是研究大量的随机现象所具有的内在规律性，即所谓统计规律性．

1.2. 古典概型(简单的概率空间)　在上面我们已经提到什么是随机现象或试验，现在再详细地讲一下，一个随机试验是指具备下列几个特点的试验．

(i) 在基本相同的条件下可重复进行多次；

(ii) 虽然不能断定某特定的结果将在一次试验中出现，但能对试验的一切可能结果进行叙述；

(iii) 在重复进行试验时，个别的结果看来是偶然发生的，

但当重复试验次数相当大时，总有某种规律性出现.

例如，掷一枚硬币，每次投掷中可能“掷出正面”，也可能“掷出反面”. 在历史上有些人对此作试验而得下列结果:

实验者	掷的次数	“掷出正面”次数	频率
Buffon	4040	2048	0.5069
K. Pearson	12000	6019	0.5016
K. Pearson	24000	12012	0.5005

从这里可看出，在相同的条件下(由同一人同一枚硬币投掷)作试验(投掷硬币)，试验的结果是“出现正面”还是“出现反面”虽不能事先断定，但我们知道试验的可能结果只有两种，最后在重复多次后，“出现正面”这个结果的相对频率却呈现出稳定性(即接近 0.5)这便是规律性.

特性(i)是我们进行试验的前提，特性(ii)无非是对试验的整体情况进行描述，即试验的一切结果都是知道的. 今后我们将试验的每个可能的结果 u 称为一个基本事件. 而将由某些基本事件所组成的结果称为事件. 而特性(iii)无非是说重复进行试验；当试验次数相当大时，某个事件的出现具有某种规律性，而这种规律性是事件本身所具有的内在属性，不随人的主观意志而变化，这种规律性可用“事件的概率”来表达.

由上可知，当我们要对一个随机现象或试验作描述时，最重要的是两件事：一是“事件”，一是“事件的概率”. 现在我们先就一种最简单的情形来讨论. 假设试验具备下列二个条件:

(i) 试验的所有可能结果的个数是有限的，设为 n，如用 u_i 表示基本事件，则试验的一切可能结果可记为 $u_1, \cdots, u_n$.

(ii) 由于具有某种对称性，认为试验的上述各种结果出现的可能性是相等的. 也就是说 $u_1, \cdots, u_n$ 是 n 个等可能

的事件. 在这样的模型中，所谓一事件便是若干个基本事件的集合. 如事件 A 由 m 个基本事件所组成（也就是说 A 中所包含的等可能的基本事件个数是 m）便定义事件 A 的概率 $P(A)$ 为:

$$P(A)=m/n. \tag{1}$$

以上的模型称为古典概型(或简单的概率空间)，而由(1)定义的 $P(A)$ 就是所谓概率的古典定义.

对于古典概型，从概率的定义可看出它有下列的几个性质:

(a) 对于任一事件 A, $0\leqslant P(A)\leqslant 1$;

(b) $P(\Omega)=1$, Ω 为必然事件;

(c) 设事件 $A_1, A_2, \cdots, A_m (m\leqslant n)$ 互不相容（即 $A_i\cap A_j=\phi, i\neq j, \phi$ 表示不可能事件），则 $P\left(\bigcup_{i=1}^{m} A_i\right)=\sum_{i=1}^{m}P(A_i)$.

事实上，(a) 由定义里就成立. (b) 因 Ω 是由全体基本事件 $u_1, \cdots, u_n$ 所组成，因而

$$P(\Omega)=\frac{n}{n}=1.$$

最后设 $A_i=(u_1^{(i)}, \cdots, u_{k_i}^{(i)})$ 是由 $k_i(\leqslant n)$ 个不同的等基本事件所组成的，那末

$$P(A_i)=\frac{k_i}{n},$$

而事件

$$\bigcup_{i=1}^{m} A_i=(u_1^{(1)}, \cdots, u_{k_1}^{(1)}; u_1^{(2)}, \cdots, u_{k_2}^{(2)}; u_1^{(m)}, \cdots, u_{k_m}^{(m)})$$

包含 $\sum_{i=1}^{m} k_i$ 个不同的基本事件，故

$$P\left(\bigcup_{i=1}^{m} A_i\right)=\frac{\sum_{i=1}^{m} k_i}{n}=\sum_{i=1}^{m}\frac{k_i}{n}=\sum_{i=1}^{m}P(A_i).$$

为了计算事件的概率，现在简略地回忆一下有关排列组合的知识.

1.3. 排列与组合

(a) 不同事物的 r-排列　由 n 个不同的事物中抽出 r 个加以排列（也就是按次序地抽出 r 个事物），每一种这样的排法都称为一个 r-排列. 它们的个数记为 $P(n, r)$ 或 ${}_nP_r$，则

$$P(n, r)=n(n-1)\cdots(n-r+1),\ r\leqslant n. \tag{2}$$

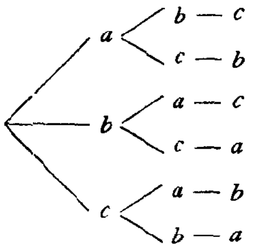
特别，

$$P(n, n)=n! \tag{3}$$

我们规定

$$P(n, 0)=0!=1.$$

公式 (2) 或 (3) 的推导由下列的图形便容易了解:



$(r=n=3$ 的情形)

利用 (3) 可把 (2) 改写成:

$$P(n, r)=\frac{n!}{(n-r)!}=\frac{P(n, n)}{P(n-r, n-r)}.$$

(b) n 个事物中，r_1 个为同一类，r_2 个又为同一类如此等等，这样的类共有 k 个，$(r_1+r_2+\cdots+r_k=n)$. 求这 n 个事物的排列数（n-排列的数）.

设要求的排列数是 x，设想把 r_1 个相同的事物换为 r_1 个新的各不相同且与其余的事物也不相同的事物，那末这 r_1 个事物可以有 $r_1!$ 种排法，这样改换后的 n 个事物便有 $x(r_1!)$ 种排法，对其余各类也应用同样的设想便得出关系式：

$$x \cdot r_1! r_2! \cdots r_k! = n!$$

于是得到公式：

$$x = n!/r_1! r_2! \cdots r_k!, \quad r_1 + r_2 + \cdots + r_k = n. \tag{4}$$

(c) 对 n 个事物允许重复抽取时的 r-排列　设有 n 个事物，每次抽取一事物后仍放回，再进行下次的抽取，对于这种情况，可求得 r-排列的个数为

$$U(n, r) = n^r. \tag{5}$$

(d) n 个不同事物的 r-组合　从 n 个不同的事物中选取 r 个(注意，这时不考虑 r 个事物的先后次序)，每一个选择所得的 r 个事物便称为一个 r-组合. 现在要求 r-组合的总数 $C(n, r)$. 只要注意到对 r 个不同的事物的一个组合把其元素排列，则有 $r!$ 种不同的排法，因此

$$r!C(n, r) = P(n, r) = n(n-1)\cdots(n-r+1),$$

$$n \geqslant r,$$

即

$$C(n, r) = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!}. \tag{6}$$

我们规定

$$C(n, 0) = 1.$$

(e) 允许重复的组合　设有 n 个不同的事物，在构成一些 r-组合时，每个事物都允许重复(从 0 次到 r 次)，现要求这样情形下的组合总数 $f(n, r)$.

把这 n 个事物记为 $1, 2, \cdots, n$. 设$(C_1, C_2, \cdots, C_r)$是任一个允许重复的组合，这个组合是按增大次序写出的，其中有

些 C_i 可能是相等的. 现按照下列方法构成一个组合 $(d_1, d_2, \cdots, d_r)$ 使

$$d_1 = C_1 + 0, d_2 = C_2 + 1, \cdots, d_i = C_i + (i-1), \\ \cdots, d_r = C_r + (r-1).$$

可知 C_i 的组合数和 d_i 的组合数是相等的, 但后者是 $1, 2, \cdots, n+r-1$ 这 $(n+r-1)$ 个不同的数的 r-组合数, 因此

$$f(n, r) = C(n+r-1, r) = \binom{n+r-1}{r}. \qquad (7)$$

1.4. 古典概率计算的例

例 1 一批产品中有 m 件一等品, n 件二等品, 从其中任抽 b 件. 问全是二等品的概率是多少?

解 任抽 b 件或说成是随机地抽取 b 件, 这表示每件产品被抽出的可能性是相等的, 这样, 所有的基本事件便是 $\{(C_1, C_2, \cdots, C_b)\}$, 其中 C_i 表示一等品或二等品. 而我们所关心的事件 A 是: b 件都是二等品, A 便是一些基本事件 $(C_{i_1}, C_{i_2}, \cdots, C_{i_b})$ 和所成的集合, 其中 $C_{i_1}, C_{i_2}, \cdots, C_{i_b}$ 都表示二等品. 于是 A 中所含的基本事件的个数是 $C(n, b)$, 而一切可能的基本事件数是 $C(m+n, b)$, 因此得

$$P(A) = \frac{C(n, b)}{C(m+n, b)}.$$

例 2 数字 $1, 2, \cdots, n$ 随机地排列, 问 2 紧接在 1 的后面的概率是多少?

解 题中所设"随机地排列"也就意味着总数为 $n!$ 的一切可能的基本事件是等可能的, 适合本题中所要求的那样一些基本事件的全体是事件 A. 有利于 A 发生的情形是: "1"有 $(n-1)$ 个位置, 而对"1"的每个位置, "2"只有一个可能的位置, 当"1"与"2"位置定了以后, 其他整数有 $(n-2)!$ 种

排法，因此有利于事件 A 发生的有 $(n-1)\cdot 1\cdot(n-2)!$ 种情形，于是得

$$P(A)=\frac{(n-1)\cdot 1\cdot(n-2)!}{n!}=\frac{1}{n}.$$

例 3 从装有 0, 1, $\cdots$, 9 十个号码的袋中抽后放回地抽出 k 个号码，问"0"恰好出现 3 次 $(3\leqslant k)$ 的概率是多少?

解 这时等可能的结果的总数是 10^k，而在 k 次抽取中"0"出现 3 次的方式有 $\binom{k}{3}$ 种，对于每种"0"出现三次的情形，其余 $(k-3)$ 次抽号的方式有 9^{k-3} 种，因此所求的概率为

$$P(A)=\binom{k}{3}\frac{9^{k-3}}{10^k}.$$

1.5. 一般的概率空间 以前所讨论的概率空间只限于基本事件的总数是有限的，而且每个基本事件都是等可能的，但是近代概率论的应用所涉及的情况却多不如此. 所以我们有必要根据古典概型所提供的有关事件与概率的一些特征性质加以抽象地概括而给出一般概率空间的定义，也就是对全部基本事件的集合，事件的集合以及事件概率给出公理化的叙述.

跟以前一样，把随机试验的一切可能的结果，都称为基本事件，记为 ω. 而把 ω 的全体记为 $\Omega=\{\omega\}$，称 Ω 为样本空间. 而由一些基本事件所组成的集合(即 Ω 的子集)便称为事件. 如基本事件 ω 属于 A 便记为 $\omega\in A$，而事件 A 出现，便是属于 A 的某个 ω 出现. 对于事件有下列一些关系:

1° Ω 是必然事件，不可能事件记为 ϕ;

2° $A\subset B$. 如事件 A 发生，必然导致事件 B 发生，则说事

件B包含事件A，记为$B\supset A$或$A\subset B$．这也就是说$\omega\in A\Longrightarrow\omega\in B$；

3° $A\cup B$．事件A与B中至少有一个发生所构成的事件称为事件A与B之和(并)，记为$A\cup B$，即，$\omega\in A$或(且)$\omega\in B\Longleftrightarrow\omega\in A\cup B$；

4° $A\cap B$．由事件A与B同时发生而构成的事件，称为它们的交，记为$A\cap B$，或简记为AB．即，$\omega\in A$且$\omega\in B\Longleftrightarrow\omega\in A\cap B$；

5° $A\cap B=\phi$，则说事件A与B互不相容，即$\omega\in A\Longrightarrow\omega\notin B$；

6° $\bar{A}$．"事件A不发生"的那个事件记为$\bar{A}$；

7° $A-B$．事件A发生而B不发生，即$\omega\in A$但$\omega\notin B$，或即$A-B=A\cap\bar{B}$；

对于样本空间Ω和它的事件族(子集族)$\mathscr{U}$我们要求满足下列条件：

(I) $\Omega\in\mathscr{U}$；

(II) 如$A\in\mathscr{U}$，则$\bar{A}\in\mathscr{U}$；

(III) 设$\{A_i\}$为一系列事件，对于一切i，$A_i\in\mathscr{U}$，则$\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{U}$．

定义1 给出集合Ω，Ω的子集族$\mathscr{U}$如满足条件(I)，(II)，(III)则称之为一个σ-代数(或σ域)．

用这个定义的术语，我们今后所讨论的概率模型中的事件便是由相应的集合Ω所得出的某个σ-代数中的子集而已．这样，以后对于事件与子集往往不加区别地通用．

现在我们可以根据古典概率或相对频率所提出的一些性质给出一般的概率定义如下．

定义2 事件的概率P是定义于σ代数$\mathscr{U}$上满足下列

三个条件的一个函数(所谓集函数).

(A) $P(A) \geqslant 0, \forall A \in \mathscr{U}$;

(B) $P(\Omega) = 1$;

(C) 设 $A_i(i = 1, 2, \cdots)$ 互不相容(即 $A_i \cap A_j = \phi$, $i \neq j$). 则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

由这个定义可以推出概率 P 的其他一些重要性质,如

1° $P(\phi)=0$. 事实上,令 $A_1 = \Omega$, $A_2 = A_3 = \cdots = \phi$, 则 $P(\Omega) = 1 = P(\Omega) + P(\phi) + P(\phi) + \cdots = 1 + P(\phi) + P(\phi) + \cdots$. 由于 $P(\phi) \geqslant 0$, 故 $P(\phi) = 0$;

2° 如 $A_1, A_2, \cdots, A_n$ 互不相容,则 $P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$. 事实上,在(C)中令 $A_{n+1} = \phi$, $A_{n+2} = \phi, \cdots$ 即可;

3° 对任两事件 A 与 B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. 因为, $A \cup B = A \cup \bar{A}B$, 故由 2°, $P(A \cup B) = P(A) + P(\bar{A}B)$, 又因 $B = AB \cup \bar{A}B$, 故 $P(B) = P(AB) + P(\bar{A}B)$, 即

$$P(\bar{A}B) = P(B) - P(AB),$$

故证得 3°;

4° $P(A \cup B \cup C) \leqslant P(A) + P(B) + P(C)$. 这是 3° 的推论;

5° $A \subset B \Longrightarrow P(A) \leqslant P(B)$, 因为, $B = BA \cup B\bar{A} = A \cup B\bar{A}$, 故 $P(B) = P(A) + P(B\bar{A})$, 但 $P(B\bar{A}) \geqslant 0$, 故证得 5°;

6° $P(\bar{A}) = 1 - P(A)$, 因为 $\Omega = A \cup \bar{A}$, 故 $1 = P(\Omega) = P(A) + P(\bar{A})$.

定义 3 由集合 Ω, Ω 的子集族所构成的一个 σ-代数 $\mathscr{U}$

以及在 $\mathscr{U}$ 上定义的某个概率函数 $P(A)$ 所成的三元组 $(\Omega, \mathscr{U}, P)$，称为一个概率空间.

例如，设集合 Ω 是可数的，$\Omega = \{\omega_1, \omega_2, \cdots\}$，又用 $\mathscr{S}(\Omega)$ 表示 Ω 的一切子集(包含 Ω 本身)所构成的族，则显见 $\mathscr{S}(\Omega)$ 是一个 σ-代数. 对于每个 ω_i 给出一个实数 $P(\omega_i)$ 与之对应，其中 $0 \leqslant P(\omega_i) \leqslant 1$，且 $\sum_{i=1}^{\infty} P(\omega_i) = 1$. 定义 $P(A) = \sum_{\omega_i \in A} P(\omega_i)$，则 $(\Omega, \mathscr{S}(\Omega), P)$ 便是一个概率空间.

习　题　1

(1) 验证递推关系式

$$P(n, r) = P(n-1, r) + rP(n-1, r-1).$$

(提示：除利用定义直接验证外，还可利用 r-排列是否包含一指定事物，并对所有 r-排列按含该指定事物分成二类的方法而推导之.)

(2) 证明

$$C(n, r) = C(n-1, r-1) + C(n-1, r),$$

并证明二项式系数间有下列关系

$$\begin{aligned} C(n,r) &= C(n-1,r-1) + C(n-2,r-1) + \cdots + C(r-1,r-1) \\ &= C(n-1,r) + C(n-2,r-1) + \cdots + C(n-1-r,0). \end{aligned}$$

在这里，当 $k < 0$ 或 $k > m$ 时规定 $C(m, k) = 0$.

(3) 设 $f(x_1, x_2)$ 的 n 阶混合导数都是连续的，求

$$\frac{\partial^n f}{\partial x_1^r \partial x_2^{n-r}}$$

的个数.

(4) 在第一组 $(a_1, \cdots, a_{n_1})$，第二组 $(b_1, \cdots, b_{n_2})$，$\cdots$ 第 k 组 $(g_1, \cdots, g_{n_k})$ 的产品中分别抽取 r_1 个，r_2 个，$\cdots$，r_k 个构成一组 $(r_i \leqslant n_i, i = 1, \cdots, k)$，这样的抽取法共有多少种?

(5) 为了研究宇宙空间，若将 6 支不同类型的火箭同时发射到 6

个不同的轨道上去，问有多少种发射方法？如不是同时发射的，而是逐支发射，又有多少种方式？又在此6支中任意取定3支而同时发射，又有多少种方式？

(6) 在上题中假设有3支火箭是同一类型的，那末将6支火箭同时发射，共有多少种方式？

(7) 有9本不同的参考书，借给一个班的3个组，每组借3本，共有多少种借法？

(8) 把1，2，3，4，5诸数字各写在一纸片上，任取三张并自左向右排列之，问所得的三位数是偶数的概率是多少？

(9) 甲、乙两乒乓球队，每队有队员6人，如混合起来分成三组进行双打表演，问每组中各有甲、乙队员两人的概率是多少？

(10) 一批产品共有n件，其中一厂生产的为n_1件，二厂生产的为n_2件，…，k厂生产的为n_k件$(n_1+n_2+\cdots+n_k=n)$。今抽取m件进行检验，问其中有m_1件为一厂生产的，m_2件为二厂生产的，…m_k件为k厂生产的概率是多少？$(m_1+m_2+\cdots+m_k=m,\quad m\leqslant n)$

(11) 有m个人站成一行，其中有二人姓张和姓王，问在张与王之间有r个人的概率是多少？

(12) 证明下列的集合等式：

(a) $(A\cup B)-B=A\cap\bar{B}$；

(b) $(A-B)\cup(B-A)=(A\cup B)-(A\cap B)$；

(c) $(A-B)\cap B=\phi$；

(d) $(A\cup B)\cap(A\cup C)=A\cup(B\cap C)$；

(e) $\bar{A}\cap(\overline{B\cap C})=(\bar{A}\cap\bar{B})\cup(\bar{A}\cap\bar{C})$；

(f) $(\overline{\bar{A}})=A$。

(13) 证明下列诸等式：

(a) $\overline{A\cap B}=\bar{A}\cup\bar{B}$，一般地，$\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\bar{A}_i$；

(b) $\overline{A\cup B}=\bar{A}\cap\bar{B}$，一般地，$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\bar{A}_i$。

(14) 从1到10中随机地选一个数字，