MENGTEKALUO FANGFA YINLUN

蒙特卡罗方法引论

朱本仁 编

山东大学出版社

蒙特卡罗方法引论

朱本仁 编

山东大学出版社

蒙特卡罗方法引论

朱 本 仁 编

山东大学出版社出版

山东省新华书店发行

山东大学印刷厂印刷

787×1092毫米 5.75 印张 129 千字

1987年8月第1版 1987年8月第1次印刷

印数 1,500册

ISBN 7-5607-0024-1/N·2

书号：13338·22 定价 0.98元

序　　言

蒙特卡罗（*Monte Carlo*）方法又名随机模拟法或统计试验法。它是在第二次世界大战期间兴起和发展起来的。它的奠基人是冯·诺伊曼（*J.Von Neŭmann*）。其主要思想是在计算机上模拟实际概率过程，然后加以统计处理。这种方法和传统数学方法相比，具有思想新颖，直观性强、简便易行的优点，它能处理一些其他方法所不能的复杂问题，并且容易在计算机上实现。从目前有关文献看，蒙特卡罗方法已经在应用物理，原子能，固体物理，化学，生物，生态学，社会学以及经济行为等领域中得到广泛的应用。特别是在计算机上用蒙特卡罗方法解很多理论和应用科学问题，在很大程度上可以代替许多大型的、难以实现的复杂实验或社会行为过程。所以，可以毫不夸张地说，由于有了蒙特卡罗方法，计算机已经不仅仅是数学家和理论科学家的重要工具，也已经成为许多实验物理家、应用科学家、社会学家的第二实验场所。

本书的对象主要是物理、化学和其他有关应用科学专业的研究生或高年级学生、专业科学工作者。他们希望通过短期的学习，能学会使用此方法来解决本专业中提出的许多实际问题。因此，本书内容的编排是根据在一学期时间内（大约50学时左右），使学生基本掌握蒙特卡罗方法的基本概念和方法这一指导思想进行的。这样的教材在国内外尚难找到，所以根据我们手头的资料决定主要取材于参考书目〔2〕

中第一至五章，以及〔4〕、〔3〕中的若干内容，还有编者个人的若干想法及推算。考虑到本书的读者一般不具备概率论的完备知识，本书根据内容的需要在第一章编写了有关概率、统计方面的最低限度的基本知识。这样使本书自成体系，读者不需要在学习中查阅其他书目和文献，就能学懂所有内容。从第二章开始，在每章末按排了少量能在小型计算机上完成的实验题。这些题目的选择既利于巩固已学的知识，也为后续章节作了准备。通过实验还能提高读者的学习兴趣。因而自然要求读者能会运用一、两种算法语言（例如 *Basic*，*Fortran*语言）和具备使用计算机的能力。为了顺利阅读本书，要求读者具备理科高等数学知识。

本书内容曾给山东大学物理系84—85学年，85—86学年的研究生以及部分教师讲授过，现经修改作为正式教材出版。在讲课和本书编写过程中承蒙物理系梅良模教授以及其他有关教师热情鼓励和大力支持，在此仅表真挚的谢意。

由于编者水平有限，错误在所难免，恳切希望读者批评指正。

朱本仁　　于山东大学

1986年1月7日

目　录

第一章　概率论基本知识

§1　事件和概率

在我们观察客观现象时，可能有三种情况发生：

（1）　在一定条件下必然会发生的某种“事件”（现象），称为必然事件。如在标准大气压下，水加热到100℃时必然会沸腾，在地球表面附近重物必然会下落；太阳必然从东边升起；……

（2）　在一定条件下必然不会发生的“事件”，称为不可能事件。如在标准大气压下，低于100℃的水不会沸腾；在地球表面附近重物无外因不会自己上升；太阳从西边升起；………

（3）　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。如打靶中环数；中彩票的事件；赌博的输赢；还有下列实例：

例1　“在一分钟内，一个电话交换台至少接到15次呼唤”或“少于15次呼唤”这类事件。

例2　“在抽查某工厂生产的10件产品时，发现有一件次品”、“发现有两件次品”、“至少发现一件次品”和“不发现次品”的事件。

例3　“从一个袋中摸黑、白球的试验中，摸出黑球”、“连续两次摸出黑球”的事件。

为了详细说明问题，我们举一个产品检查的具体实例来讨论。

例 4　在一大批产品中分别抽 5 件、10件、60件、150件、600件、900件、1200件、1800件产品来检查，结果记录如下：

表1—1

抽取件数	5	10	60	150	600	900	1200	1800
合格品数	5	7	53	131	548	820	1091	1631
合格品频率	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

$$合格品频率=\frac{合格数}{抽取件数}$$

虽然抽出的产品中，次品数目是随机的，然而随着抽查件数的增加，合格品频率愈来愈趋于一个稳定値0.9。

现在引进概率的定义。设U为必然事件，V为不可能事件，A为一般的随机事件，它们的概率是一非负实数，分别记之为$P(U)$，$P(V)$和$P(A)$，且满足：

$$P(U)=1,\ P(V)=0，通常P(A)满足$$

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

$P(A)$表示出现事件A的可能性的大小，将其称为事件 A 的**概率**。

事件A的概率$P(A)$可理解为事件 A 的函数。今后，我们用大写字母A，B，C，…表示事件，而用$P(A)$，$P(B)$，$P(C)$，…表示相应事件的概率。事件间主要有下列几种关系：

（1）　如果事件A发生，必然导致事件B发生，则称

事件A是事件B的**特款**，记作$A\subset B$或$B\supset A$。

若$B\supset A$且$A\supset B$，则称A与B**等价**，记作$A=B$.

（2） 事件A或B发生所构成的事件，称为事件A与B的**和**。记作$A\cup B$。其推广为事件A_1或A_2或……或A_N发生所构成的事件，称为事件A_1，A_2，……A_N的和，记作$\bigcup\limits_{n=1}^{N}A_n$。

（3） 事件A或B同时发生所构成的事件，称为事件A与B的**交**，记作$A\cap B$（简记为AB）。其推广为事件A_1，A_2，……，A_N同时发生所构成的事件，记为$\bigcap\limits_{n=1}^{N}A_n$。

（4） 若事件A发生，必然导致B不发生（反之亦然），则称A与B是**互不相容**的事件。即有$AB=V$。

若记$\bar{A}$表示事件A不发生的事件，称A与$\bar{A}$为**互逆**事件。于是，显然有

$$A\cup\bar{A}=U,\qquad A\cap\bar{A}=V.$$

事件A的概率$P(A)$的基本性质（见图1—1，a、b、c、d）：

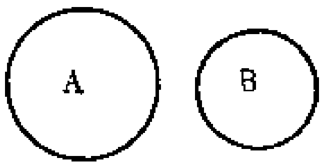

图1—1，a

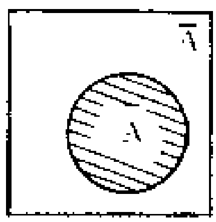

图1—1，b

（1） 若$A\cap B=V$，则$P(A\cup B)=P(A)+P(B)$.

（2） 若$\bar{A}$为A的逆事件，则

$$P(\bar{A})=1-P(A)$$

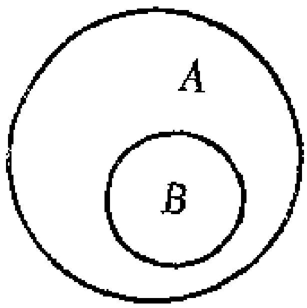

图1—1，c

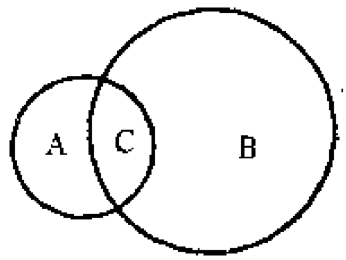

图1—1，d

（3） 若$A\supset B$，则$P(A)\geqslant P(B)$

（4） 若$A\cap B=C\neq V$，则

$$A=A\cap(B\cup\bar{B})=\underline{A\cap B}\cup\underline{A\cap\bar{B}},$$

$$B=B\cap(A\cup\bar{A})=\underline{B\cap A}\cup\underline{B\cap\bar{A}},$$

$$A\cup B=\underline{A\cap\bar{B}}\cup\underline{A\cap B}\cup\underline{B\cap\bar{A}},$$

$$P(A\cup B)\leqslant P(A)+P(B)。$$

现在引进**条件概率**：$P(A|B)$表示在“事件B发生的条件下”事件A发生的概率。考虑事件集合B和A（见图1—1，e）

$$P(A|B)=\frac{P(AB)}{P(B)}$$

由此可知**乘法定理**：

$$P(AB)=P(A|B)P(B)$$
$$=P(B|A)P(A).$$

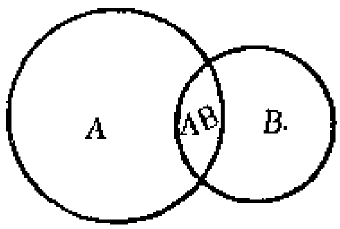

图1—1，e

例5 设96件产品中有5件次品，任意抽查两件，求两件都合格的概率

解：以A代表第一次抽得合格品的事件，以B代表第二次抽得合格品的事件。由乘法定理

$$P(AB)=P(A)\cdot P(B|A)$$

而 $$P(A)=91/96,\quad P(B|A)=90/95$$

所以 $$P(AB)=\frac{91\times 90}{96\times 95}$$

$P(AB)$就是所抽的两件都是合格品的概率。由于

$$B=B(A\cup\bar{A})=BA\cup B\bar{A}$$
$$P(B)=P(BA)+P(B\bar{A})$$
$$P(B\bar{A})=P(\bar{A})P(B|\bar{A})$$
$$P(\bar{A})=5/96,\qquad P(B|\bar{A})=91/95$$

所以

$$P(B)=\frac{90\times 91}{96\times 95}+\frac{5\times 91}{96\times 95}=\frac{91}{96}\neq P(B|A)$$

这说明了由于事件A的出现影响了事件B的概率，表明$P(B|A)$与A有关。但有时若事件A不影响事件B的概率，即

$$P(B|A)=P(B)\text{或}P(AB)=P(A)P(B)$$

则称事件A与B是**独立的**。

例 6　由两道工序加工圆柱体，直径不合格率为p，长度不合格率为q。A表示直径不合格的事件，B表示长度不合格的事件。若该两道加工工序完全无关，那末应有

$$P(A)=p,\quad P(B)=q,\quad P(AB)=P(A)P(B)=pq$$

这表明事件A和B是独立的。

§2 概　率　场

在考察任一随机现象时，每次的试验或观察总可得一结果，称此结果为**基本事件**，基本事件的全体为**基本空间**，记为U。若把基本空间视为一个集合，其元素为基本事件，

在问题的研究中还要涉及U中的某些子集合A，B，……这些子集合的全体记为$\mathscr{F}$，要求$\mathscr{F}$满足：

（1） $U\in\mathscr{F}$；

（2） 若$A_n\in\mathscr{F}$，$n=1$，2，…，则$\bigcup_n A_n\in\mathscr{F}$；

（3） 若$A\in\mathscr{F}$，则$\bar{A}\in\mathscr{F}$。

由上面（1），（3）立即知$V\in\mathscr{F}$。

在$\mathscr{F}$上定义概率P：任取$A\in\mathscr{F}$，$P(A)$应满足：

（1） $0\leqslant P(A)\leqslant 1$， $A\in\mathscr{F}$；

（2） $P(U)=1$；

（3） 若$A_n\in\mathscr{F}$，$A_i\cap A_j=V$，$i\neq j$， 则：

$P(\bigcup_n A_n)=\sum_n P(A_n)$，称为概率的完全可加性。

这样的U，$\mathscr{F}$，P称之为一个**概率场**，记作（U，$\mathscr{F}$，P），有时也称它为一个**概率空间**。

根据上述公理，我们立即可以推出一些重要的结论：

（1） 不可能事件的概率为零，即$P(V)=0$。

（2） 若$A_i\cap A_j=V(i\neq j)$，则

$$P(\bigcup_{j=1}^{n}A_j)=\sum_{i=1}^{n}P(A_i)$$

（见图1—2，a）

（3） 任取$A\in\mathscr{F}$，有$P(\bar{A})=1-P(A)$.

（4） 若$A\supset B$（见图1—2，b），则

$$P(A-B)=P(A)-P(B)$$

（5） 若$A\supset B$，则$P(A)\geqslant P(B)$.

（6） 一般情况（见图1—2，c）我们有

$$P(A\cup B)\leqslant P(A)+P(B),$$

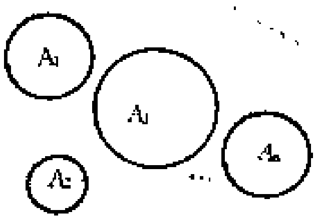

图1—2，a

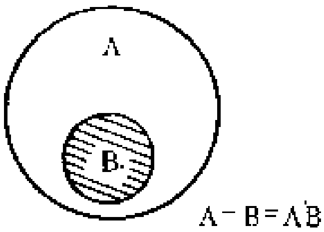

图1—2，b

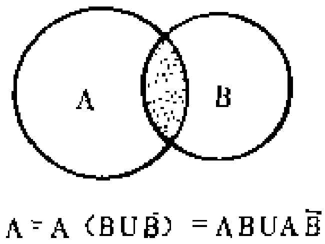

图1—2，c

$$P(\bigcup_i A_i)\leqslant \sum_i P(A_i).$$

在概率场$(U,\mathscr{F},P)$中定义条件概率$P(A|B)$如下：

$$P(A|B)=P(AB)/P(B)\qquad (P(B)\neq 0).$$

若“已知事件B发生”的条件对事件A的概率不发生影响，即有$P(A|B)=P(A)$，此时有

$$P(AB)=P(A)P(B)$$

称**事件A对于事件B独立。**

由条件概率及事件独立性的定义，可以推出下列结果。

（1） 若事件A对于事件B独立，$P(A)\neq 0$，$P(B)\neq 0$，则事件B对于事件A也独立，即两个事件的独立性是对称的。因为

$$P(AB)=P(A|B)P(B)=P(B|A)P(A)$$
$$=P(A)P(B).$$

我们称为$\{A,B\}$独立。

（2） 若$\{A,B\}$独立，则$\{\bar{A},B\}$，$\{A,\bar{B}\}$，$\{\bar{A},\bar{B}\}$亦都独立。因为

$$\bar{A}B=B-AB,$$

$$P(\bar{A}B)=P(B)-P(AB)=P(B)-P(A)P(B)$$
$$=P(B)(1-P(A))=P(B)P(\bar{A}).$$

独立性的概念可以推广至多个事件的情况：一组事件 $A_1, A_2, \cdots, A_s$称为**相互独立的**，那么对任何$r(1\leqslant r\leqslant s)$及$1\leqslant i_1<i_2<\cdots<i_r\leqslant s$有

$$P(A_{i_1}A_{i_2}\cdots A_{i_r})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_r}).$$

应当注意，两两独立的一组事件不一定是相互独立的。例如一个均匀的四面体，第一面染红色记为A，第二面是绿色记为B，第三面是蓝色记为C，第四面同时染有红、绿、蓝色记为(ABC)。投掷四面体时出现A, B, C的概率均为1/2(见图1—2，d)

$$P(A)=P(B)=P(C)=\tfrac{1}{2}，而$$

$$P(AB)=P(BC)=P(CA)=\tfrac{1}{4},$$

然而 $P(ABC)=\frac{1}{4}\neq P(A)P(B)P(C)$.

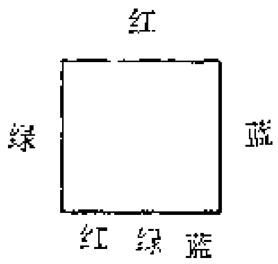

图1—2，d

（3） **全概率公式** 设事件B能而且只能与互不相容的事件$A_1, A_2, \cdots, A_n, \cdots$之一同时发生，即

$$B=\bigcup_{i=1}^{\infty}BA_i$$

这里当$i\neq j$时由于$A_iA_j=V$，因而$(BA_i)(BA_j)=V$，表明BA_i与BA_j是互不相容的，由概率的完全可加性有

$$P(B)=\sum_{i=1}^{\infty}P(BA_i)$$

利用 $P(B|A_i)=P(BA_i)/P(A_i)$

就可推得 $P(B)=\sum_{i=1}^{\infty}P(A_i)P(B|A_i)$

这就是全概率公式。

全概率公式的几何意义见图1—2，e：B的面积为$P(B)$，A_i的面积为$P(A_i)$，B在A_i中所占有的面积与A_i的面积的比值为$P(B|A_i)$。

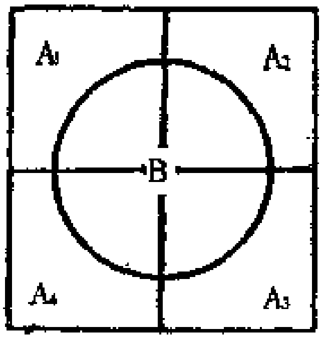

图1—2，e

（4） **贝叶斯(Bayes)公式** 设事件B能而且只能与互不相容的事件A_1，A_2，…，A_n，…之一同时发生，即

$$B = \bigcup_{i=1}^{\infty} BA_i$$

于是从 $P(A_iB) = P(B)P(A_i|B) = P(A_i)P(B|A_i)$

得 $P(A_i|B) = P(A_i)P(B|A_i)/P(B)$

利用全概率公式

$$P(A_i|B) = P(A_i)P(B|A_i)/\sum_{j=1}^{\infty} P(A_j)P(B/A_j)$$

这就是贝叶斯公式。

例1　在一段时间t内某电话站得到K次呼唤的事件A_t^K的概率为$P_K(t)$，假设在同一时间长度内得到K次呼唤的概率相同，且两段时间内各自得到呼唤的事件是相互独立的，求在$2t$时间内得到S次呼唤的概率。

解：　$A_{2t}^{s} = A_t^0 A_t^s \cup A_t^1 A_t^{s-1} \cup \cdots \cup A_t^s A_t^0$

$$P(A_{2t}^{s}) = P_s(2t) = \sum_{i=0}^{s} P(A_t^i A_t^{s-i})$$

$$= \sum_{i=0}^{s} P(A_t^i)P(A_t^{s-i}) = \sum_{i=0}^{s} P_i(t)P_{s-i}(t)$$

已知 $P_K(t)=\dfrac{(at)^K}{K!}e^{-at}$， 这里 $a=const$.

于是

$$P(A_{2t}^s)=P_s(2t)=\sum_{i=0}^{s}\frac{(at)^s}{i!(s-i)!}e^{-2at}$$

$$=\frac{(at)^s}{s!}\cdot e^{-2at}\sum_{i=0}^{s}\frac{s!}{i!(s-i)!}$$

但 $$\sum_{i=0}^{s}\frac{s!}{i!(s-i)!}=(1+1)^s=2^s$$

所以 $$P_s(2t)=\frac{(2at)^s}{s!}e^{-2at}$$

例 2 带有吸收壁的随机游动。一个质点从 x 轴的某一点 K 出发（$0<K<N$，且 K，N 均为整数），随机地在直线上游动。每次向右移动一个单位的概率为 p，向左移动一个单位的概率为 $q=1-p$．当质点游动到位置0和 N 时游动就停止，即质点被吸收。0和 N 称为吸收壁。

设 P_K^0 和 P_K^N 分别表示质点从 K 点出发而被吸收壁 0 和 N 吸收的概率。求 P_K^0 和 P_K^N.

解：设 B 表示从 K 点出发被0吸收的事件，B_+ 表示从 $K+1$ 点出发被0吸收的事件，B_- 表示从 $K-1$ 点出发被0吸收的事件，则有

$$B=B\cap\{K\to K+1\}\cup B\cap\{K\to K-1\}$$
$$=B_+\cap\{K\to K+1\}\cup B_-\cap\{K\to K-1\}$$

其中 $\{K\to K+1\}$ 表示质点从 K 点游动到 $K+1$ 点的事件，

因而

$$P_K^0 = pP_{K+1}^0 + qP_{K-1}^0 \qquad K \neq 0,\ N$$

$$P_0^0 = 1,\quad P_N^0 = 0,\ p = P(K+1\,|\,K),\ q = P(K-1\,|\,K)$$

上述方程的特征方程为

$$p\lambda^2 - \lambda + q = 0$$

有两根 $\lambda = 1,\ q/p$

其通解为 $P_K^0 = C_0 + C_1(q/p)^K$

选择C_0，C_1使之满足端点条件后解得

$$P_K^0 = \frac{q^N - q^K p^{N-K}}{q^N - p^N} \qquad K = 0, 1, 2, \cdots, N$$

而 $P_K^N = 1 - P_K^0$.

例 3 在炮战中，分别在距目标250米，200 米， 150米处射击的概率为0.1，0.7，0.2，而在各处命中的概率分别为0.05，0.1，0.2。现在已知目标被击毁，求击毁目标的炮弹是由距目标200米射出的概率。

解：设A_1为距目标250米时射击的事件，

A_2为距目标200米时射击的事件，

A_3为距目标150[illegible]

B为命中目标这[illegible]

因为 $A_1 \cup A_2 \cup A_3 = U$，因[illegible]

$$P(A_2|B) = P(A_2)\text{[illegible]}P(B|A_i)$$

$$= 0.7\times0.1/(0.05\times0.1+0.1\times0.7+0.2\times0.2) = 0.6087.$$

§3 独立试验序列概型

本节我们介绍一个基本的概型——独立试验序列概型。

首先说明“独立试验序列”的含义。一个试验序列，其各次试验的结果和其余各次试验无关（独立性）。设每次试验中所有可能出现的试验结果的全体为 A_1，A_2，…，A_K，…，其出现的概率分别为$p_1, p_2, \cdots$，$p_K\cdots$。若令第一次试验结果为 A_{j_1}，…，第 n 次试验结果为 A_{j_n}，则 n 次试验独立的含义是：

$$P\{A_{j_1} A_{j_2} \cdots A_{j_n}\} = p_{j_1} p_{j_2} \cdots p_{j_n}.$$

独立试验序列概型最简单的情况是试验结果只有两种可能A、$\bar{A}$，其出现的概率分别为p、$q=1-p$，这种情况称为贝努里（*Bernoulli*）概型。下面我们讨论贝努里概型。使我们感兴趣的是在 n 次试验中事件 A 出现 m 次（$m=0$，1，2，…，n）的概率$P_n\{m\}$。为计算它，首先注意到事件A在指定的m次试验中（例如在指定的第 2、4、6、…$2m$次（$2m \leqslant n$）试验中）发生，而在其余$n-m$次试验中不发生的概率为

由于我们未指定事件A在哪m次试验中发生，因此这里有$C_n^m=\dfrac{n!}{m!(n-m)!}$个指定位置，并且他们是不相容的，根据加法定理有

$$P_n\{m\} = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m},$$

并且显然有

$$\sum_{m=0}^{n} P_n\{m\} = \sum_{m=0}^{n} C_n^m p^m q^{n-m} = (p+q)^n = 1 \quad .$$

例 1 设每次射击打中目标的概率 p 为 0.001，如果射击5000次，试求击中两次以上的概率$P\{m\geqslant 2\}$.

解：利用 $P_{5000}\{0\}=(1-0.001)^{5000}=0.0071$

$P_{5000}\{1\}=5000\times p\times q^{4999}=0.0354$

因而有

$$P\{m\geqslant 2\} = \sum_{m=2}^{5000} P_{5000}\{m\} = 1 - P_{5000}\{0\} - P_{n5000}\{1\}$$

$$= 0.9575。$$

一般，在实际问题中利用原始表达式计算 $P_n\{m\}$是十分困难的，甚至不可能实现，尤其当n，m很大时，这个矛盾尤为突出。为此，我们给出德莫佛——拉普拉斯（*Demoivre-Laplace*）局部极限定理，该定理提供了计算$P_n\{m\}$的近似公式。

德莫佛—拉普拉斯局部极限定理 对于事件A出现的概率为$p(0<p<1)$，试验次数为n的贝努里概型，令K代表n次试验中事件A出现的次数，则对任意区间$[a, b]$，当 $a\leqslant x_K=\frac{K-np}{\sqrt{npq}}\leqslant b$及$n\to\infty$时，一致地有

$$P_n\{K\} \approx \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_K^2} \quad .$$

证明：因为$x_K\in[a, b]$，当$n\to\infty$时

$$K = np + x_K\sqrt{pnq} \to \infty$$

$$j = n - K = nq - x_K\sqrt{npq} \to \infty,$$

另一方面，根据微积分学中著名的斯梯林（Stirling）公式

$$m\,! = \sqrt{2\pi m}\, m^m e^{-m} e^{\theta} \quad (0 < \theta_m < \frac{1}{12m})$$

于是

$$P_n\{K\} = \frac{n!}{K!\,j!} p^K q^j = \frac{\sqrt{2\pi n}\; n^n e^{-n}}{\sqrt{2\pi K}\, K^K e^{-K} \cdot \sqrt{2\pi j}\; j^j e^{-j}} \cdot$$

$$\cdot\, p^K q^j e^{\theta_n - \theta_K - \theta_j} = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{Kj}} \left(\frac{np}{K}\right)^K \left(\frac{nq}{j}\right)^j e^{\theta},$$

当$a \leqslant x_K \leqslant b$时，一致有

$$|\theta| < \frac{1}{12}\left(\frac{1}{n} + \frac{1}{K} + \frac{1}{j}\right),$$

因为 $$\frac{Kj}{n} = n\left(p + x_K\sqrt{\frac{pq}{n}}\right)\left(q - x_K\sqrt{\frac{pq}{n}}\right)$$

$$= npq + (q-p)\, x_K\sqrt{npq} + \cdots \approx npq$$

$$\ln\left(\frac{np}{K}\right)^K = K\ln\left(\frac{np}{np + x_K\sqrt{npq}}\right)$$

$$= K\ln\left(\frac{1}{1 + x_K\sqrt{\dfrac{q}{np}}}\right) = -K\ln\left(1 + x_K\sqrt{\frac{q}{np}}\right)$$

$$= -\left(np + x_K\sqrt{npq}\right)\left(x_K\sqrt{\frac{q}{np}} - \frac{q}{2np} x_K^2 + \right.$$

$$\left. + O\left(\frac{1}{n^{3/2}}\right)\right) = -\left(x_K\sqrt{npq} + \right.$$

$$+\frac{q}{2}x_K^2+O(\frac{1}{\sqrt{n}}))$$

（这里用到了展开式$\ln(1+\delta)=\delta-\frac{1}{2}\delta^2+\frac{1}{3}\delta^3+\cdots$）

$$\ln(\frac{nq}{j})^j=j\ln\left(\frac{nq}{nq-x_K\sqrt{npq}}\right)$$

$$=-j\ln(1-x_K\sqrt{\frac{p}{nq}})=(nq-x_K\sqrt{npq})$$

$$(x_K\sqrt{\frac{p}{nq}}+\frac{p}{2nq}x_K^2+O(\frac{1}{n^{3/2}}))$$

$$=x_K\sqrt{npq}-\frac{p}{2}x_K^2+O(\frac{1}{\sqrt{n}}).$$

所以 $$\ln(\frac{np}{K})^K(\frac{nq}{j})^j=-\frac{x_K^2}{2}+O(\frac{1}{\sqrt{n}})$$

即 $$(\frac{np}{K})^K(\frac{nq}{j})^j=e^{-\frac{x_K^2}{2}+O(\frac{1}{\sqrt{n}})}$$

所以有

$$P_n\{K\}=\frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{Kj}}(\frac{np}{K})^K(\frac{nq}{K})^j e^\theta$$

$$=\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}(1+O(\frac{1}{\sqrt{n}}))}e^{-\frac{x_K^2}{2}+O(\frac{1}{\sqrt{n}})}\cdot e^\theta$$

$$\approx\frac{1}{\sqrt{npq}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x_K^2}$$ 证毕。

这里$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$是众所周知的正态分布函数（误差函数，见图1—3，a），它的函数值及其定积分值可以查表计算。

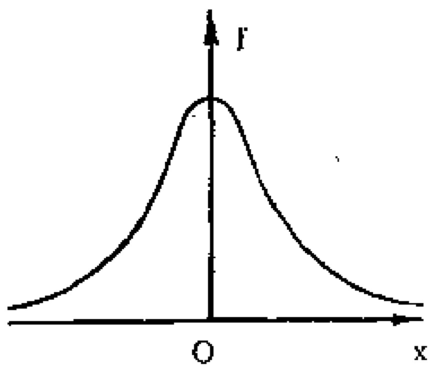

图1—3，a

为了计算事件A出现次数从K_1至K_2($K_2>K_1$)的概率，即计算$P_n\{K_1\leqslant K\leqslant K_2\}$，我们还要利用德莫佛—拉普拉斯积分极限定理。

德莫佛—拉普拉斯积分极限定理 对事件A出现概率为$p(0<p<1)$的贝努里概型，令μ代表n次试验中事件A出现的次数，则对任意指定的区间$[a, b)$恒有

$$P\{a\leqslant\frac{\mu-np}{\sqrt{npq}}<b\}\longrightarrow\int_a^b\varphi(x)dx \quad (n\to\infty)$$

其中 $$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \quad (-\infty<x<+\infty)$$

证明：$P\{a\leqslant\frac{\mu-np}{\sqrt{npq}}<b\}=P\{np+a\sqrt{npq}\leqslant\mu\leqslant np$

$$+b\sqrt{npq}\}=\sum_{K=K_1}^{K_2}P_n\{K\}$$

其中K_1为大于或等于$np+a\sqrt{npq}$的最小整数，K_2为小于$np+b\sqrt{npq}$的最大整数。由德莫佛—拉普拉斯局部极限定理，对任给的$\varepsilon>0$，当n充分大时有

$$P_n\{K\}=\frac{1}{\sqrt{npq}}(\varphi(x_K)+\varepsilon_K) \qquad |\varepsilon_K|<\varepsilon,$$

$$K\in[K_1,\ K_2]$$

代入上式得

$$P\{\ a\leqslant\frac{\mu-np}{\sqrt{npq}}<b\}=\sum_{K_1}^{K_2}\frac{1}{\sqrt{npq}}\varphi(x_K)$$

$$+\sum_{K_1}^{K_2}\frac{\varepsilon_K}{\sqrt{npq}}\ .$$

$$\left|\sum_{K_1}^{K_2}\frac{\varepsilon_K}{\sqrt{npq}}\right|\leqslant\frac{1}{\sqrt{npq}}(K_2-K_1+1)\varepsilon$$

$$\leqslant(b-a)\varepsilon$$

而当 $n\to\infty$ 时有

$$\sum_{K_1}^{K_2}\frac{1}{\sqrt{npq}}\varphi(x_K)\to\int_a^b\varphi(x)dx$$

于是得

$$P\{a\leqslant\frac{\mu-np}{\sqrt{npq}}<b\}\longrightarrow\int_a^b\varphi(x)dx \qquad 证毕.$$

例 2　设生产过程中出现次品的概率为 0.005，求在任意的10000件产品中，（1）有40件次品的概率；（2）次品不多于70件的概率。

解：(1)　$P_{10000}\{40\}=C_{10000}^{40}(0.995)^{9960}(0.005)^{40}$

$$\approx\frac{1}{\sqrt{2\pi\times10000\times0.995\times0.005}}e^{-\frac{1}{2}x_K^2}$$

其中 $$x_K=\frac{40-np}{\sqrt{npq}}=\frac{40-50}{\sqrt{50\times 0.995}}\approx\frac{-10}{7.05}=-1.42\quad,$$

所以 $$P_{10000}\{40\}\approx\frac{1}{7.05\sqrt{2\pi}}e^{-\frac{(1.42)^2}{2}}$$

$$\approx 0.0206\quad(\text{精确值为}0.0214).$$

(2) $$a=\frac{0-np}{\sqrt{npq}}\approx\frac{-50}{7.05}\approx-7.09$$

$$b=\frac{70-np}{\sqrt{npq}}\approx\frac{20}{7.05}\approx 2.84$$

$$P_{10000}\{\mu\leqslant 70\}\approx\int_{-7.09}^{2.84}\varphi(x)dx$$

令 $\Phi(x)=\int_{-\infty}^{x}\varphi(x)dx$，可以查表求得

$$\int_{-7.09}^{2.84}\varphi(x)dx=\Phi(2.84)-\Phi(-7.09)$$

$$=0.9975\quad.$$

§4 随机变量及分布函数

在本节中我们主要研究随机现象中一种重要的概型，即每次试验的结果可以用一个数ξ来表示的情形。这个变量ξ是随着试验的结果不同而变化的，我们称它为**随机变量**。换

句话说，就是用一个数ξ来代表事件，随机事件用一个随机变量来代表。这在许多实际问题中是可行的，因为我们并不关心事件本身的具体细节，例如钱币的正、反面，摸黑、白球等等，我们关心的是某一事件是否出现及其出现的概率。用随机变量的概型可以概括一大类实际生活中具体内容完全不同的随机现象。如果随机变量ξ所取的可能值能够一一地列举出来，而且ξ以各种确定的概率取这些不同的值，我们便称ξ为**离散型随机变量**（简称ξ是离散的随机变量）。

例如丢钱币的游戏中，出现钱币正面用$\xi=1$表示，出现钱币的反面用$\xi=0$表示。设出现正面的概率为p，那来出现反面的概率为$q=1-p$. 又如单位时间内电话站接到K次呼唤的事件用$\xi=K$表示，那来，我们在前面已经提到“$\xi=K$”的概率为$\frac{\lambda^K e^{-\lambda}}{K!}$。“$n$次射击中命中$K$次的事件”用$\xi=K$表示，如此等等。我们可以用下面的方式来描述这种离散的概型。

设$\{x_i|i=1, 2, \cdots\}$是离散随机变量ξ的所有可能的值，而$\{p_i|i=1, 2, \cdots\}$是ξ取这些值的相应的概率，即

$$P\{\xi=x_i\}=p_i \quad i=1, 2, \cdots (\sum_i p_i=1)$$

于是我们可用下列矩阵（分布律）来描述离散随机变量ξ：

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

在实际生活中还有一些随机现象是“连续的”。例如测量某一工件的长度时，测得的值由于各种复杂的因素可认为是随机变量，它可能取某一范围内的任何值l $(a\leqslant l\leqslant b)$。又

例如射击命中点与靶心的距离d（见图1—4，a）。在这样一些问题中关于“随机变量取某一固定值的概率”往往是无意义的。如所谓“命中环数”，实际上并非真正按打中靶子上的那一点，而是按打在靶上某一小园环区域内计算的。所以对于可以连续地从某一范围内取值的随机量，须要按一种适当方式给出它的定义：

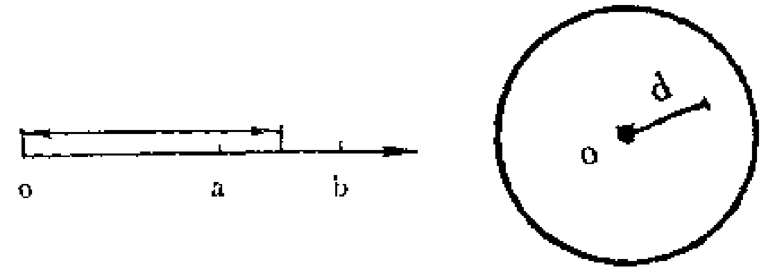

图1—4，a

定义　如果每次试验的结果，可以用一个数ξ来表示，而且对任何实数x，“$\xi<x$”有着确定的概率，则称ξ是**随机变量**。而事件$\{\xi<x\}$的概率作为x的函数

$$F(x)=P\{\xi<x\}$$

被称为随机变量ξ的**分布函数**。

现在考察事件$\{a\leqslant\xi<b\}$，在$a<b$的条件下，因为

事件$\{\xi<a\}\subset$事件$\{\xi<b\}$

事件$\{a\leqslant\xi<b\}=$事件$\{\xi<b\}-$事件$\{\xi<a\}$

事件$\{\xi<+\infty\}=U$，事件$\{\xi<-\infty\}=V$，

所以可以断定

$$P\{a\leqslant\xi<b\}=P\{\xi<b\}-P\{\xi<a\}$$
$$=F(b)-F(a)\geqslant 0$$
$$P\{\xi<+\infty\}=1,\ P\{\xi<-\infty\}=0,$$

这表明**分布**函数$F(x)$是非降的有界变差函数。对随机变量的

上述定义，它既可以描述取连续值的随机变量，又可以描述离散型的随机变量。

例 1　考虑贝努里概型。设在每次试验中，事件A出现的概率为$P(A)=p$。记μ为n次试验中事件A出现的次数，它是一个随机变量，其可能值为$\{0,1,2,\cdots,n\}$.

$$P(\mu=K)=C_n^K p^K q^{n-K}\quad(q=1-p,\ K=0,1,2,\cdots,n)$$

$$F(x)=P(\mu<x)=\sum_{K<x}C_n^K p^K q^{n-K}$$

μ的分布函数$F(x)$的图形见图1—4，b所示。

例 2　在(a,b)上任意掷一个质点，用ξ表示质点落点与原点的距离，ξ为一随机变量。若这个质点落在(a,b)中任何区间内的概率只与区间的长度成正比，则称随机变量ξ是**均匀分布**的。ξ的分布函数为

$$F(x)=\begin{cases}0 & (x<a),\\ \dfrac{x-a}{b-a} & (a\leqslant x<b),\\ 1 & (x\geqslant b).\end{cases}$$

它的图形见图1—4，c。

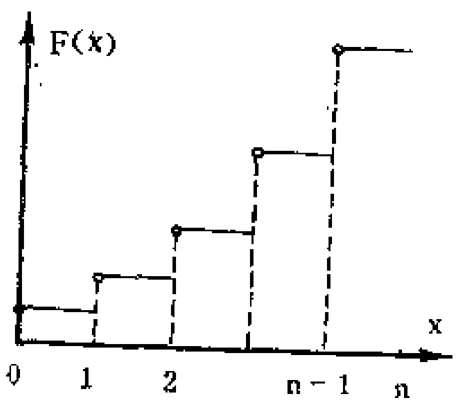

图1—4，b

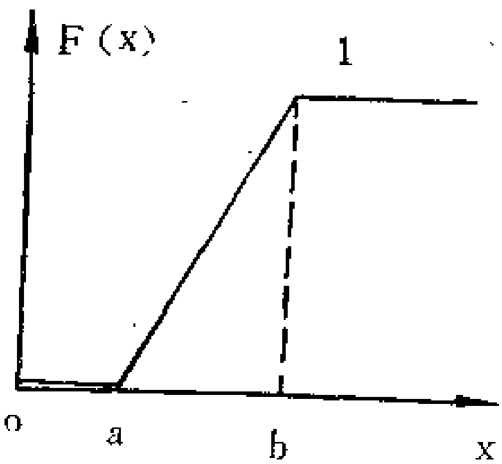

图1—4，c

定义 如果随机变量ξ的分布函数 $F(x)=P\{\xi<x\}$ 可以写成

$$F(x)=\int_{-\infty}^{x} f(z)dz$$

其中$f(x)\geqslant 0$，则称ξ为**连续型的随机变量**，而$f(x)$称为ξ的**分布密度**（或密度函数）。

在例 2 中随机变量ξ是连续型的，其次容易看出

$$f(x)=F'(x)=\begin{cases} 0 & (x\leqslant a) \\ \dfrac{1}{b-a} & (a<x<b) \\ 0 & (x\geqslant b). \end{cases}$$

容易证明密度函数$f(x)$具有下列性质：

(1) $f(x)\geqslant 0$；

(2) $\int_{-\infty}^{\infty} f(x)dx=1$；

(3) $P\{a\leqslant\xi<b\}=F(b)-F(a)=\int_{a}^{b} f(x)dx$.

例 3 对某一长度a进行测量时会发生误差，所以测得的值ξ可以看作是随机的，它服从正态分布律：

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(z-a)^2}{2\sigma^2}}dz$$

这里$\sigma>0$是一常数，它反映测量精度和可靠性。这是一个连续型随机变量，其密度函数为

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

$f(x)$的图形见图1—4，d。

最后我们指出，对一般的随机变量 ξ 的分布函数 $F(x)$而言，$F(x)$是左连续的，即 $F(x)=F(x-0)=\operatorname*{lin}\limits_{x_n\to x^-}F(x_n)$，其中序列$\{x_n\}$单调上升收敛到$x$。事实上，由于$x_1<x_2<\cdots<x_n<\cdots$。

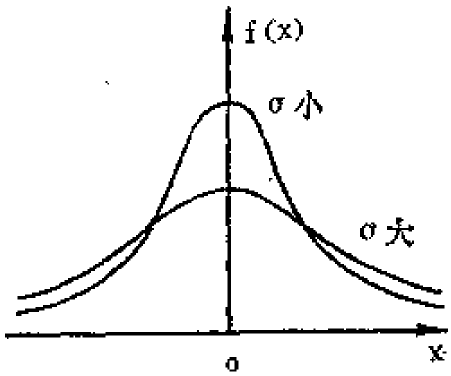

图1—4，d

$$\text{事件}\{\xi<x\}=\bigcup_n\text{事件}\{\xi<x_n\}=\bigcup_n\text{事件}\{x_{n-1}\leqslant\xi<x_n\},$$

所以

$$F(x)=P\{\xi<x\}=\sum_n P\{x_{n-1}\leqslant\xi<x_n\}$$

$$=\lim_{n\to\infty}F(x_n)。$$

§5 多元随机变量及其分布函数

在实际生活中，某些随机事件仅用一个随机变量来描述是不够的，需要同时用多个随机变量来描述。例如射击问题通常用两个随机变量$(\xi_1,\ \xi_2)$来描述更为确切。在一般情况下，我们称 n 个随机变量 $\xi_1,\ \xi_2,\ \cdots,\ \xi_n$ 的总体 $\xi=(\xi_1,\ \xi_2,\ \cdots,\ \xi_n)$为**n元随机变量**。将一元随机变量的分布函数的概念推广至 n 元的情况。称函数

$$F(x_1,\ x_2,\ \cdots,\ x_n)=P\{\xi_1<x_1,$$

$$\xi_2 < x_2,\ \cdots,\ \xi_n < x_n\}$$

为 n 个随机变量 ξ_1，ξ_2，…，ξ_n 的**联合分布函数**（n 元 分布函数）。

与一元的情况类似，容易证明 n 元分布函数 $F(x_1, x_2, \cdots, x_n)$ 具有下列三个性质：

(1) $F(x_1, x_2, \cdots, x_n)$ 对每个变元非降；

(2) $F(x_1, x_2, \cdots, x_n)$ 对每个变元左连续；

(3) $\lim\limits_{x_K \to -\infty} F(x_1, x_2, \cdots, x_n) = 0$,

$$(K = 1,\ 2,\ \cdots,\ n)$$

$$\lim_{\substack{x_1 \to +\infty \\ x_2 \to +\infty \\ \vdots \\ x_n \to +\infty}} F(x_1, x_2, \cdots, x_n) = 1;$$

$$\text{即}\quad F(+\infty, +\infty, \cdots, +\infty) = 1$$

还可将分布密度（密度函数）推广至多元的情况。若分布函数 $F(x_1, x_2, \cdots, x_n)$ 可以表为下述积分的形式

$$F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(z_1, z_2, \cdots, z_n) dz_1 dz_2 \cdots dz_n$$

则称 $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$ 为 n 元连续型随机变量（**n 元连 续随机变量**），函数 $f(x_1, x_2, \cdots, x_n)$ 为联合分布密度（密度函数）。显然，若 F 是光滑可微函数，则

$$f(x_1, x_2, \cdots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F(x_1, x_2, \cdots, x_n)$$

分布密度函数 $f(x_1, x_2, \cdots, x_n)$ 满足下列性质：

(1) $f(x_1, x_2, \cdots, x_n) \geqslant 0$;

(2) $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(z_1, z_2, \cdots, z_n) dz_1 \cdots dz_n = 1$。

多元随机变量描述的是几个随机变量的总体，因此有必要研究各个变量的**独立性问题**。n元随机变量$(\xi_1, \xi_2, \cdots, \xi_n)$的分量 ξ_1，ξ_2，…，ξ_n 是相互独立的定义：若对任意的$r(1 \leqslant r \leqslant n)$及任何$1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$，皆有

$$P(\xi_{i_1} < x_{i_1}, \xi_{i_2} < x_{i_2}, \cdots, \xi_{i_r} < x_{i_r}$$
$$= P\{\xi_{i_1} < x_{i_1}\} \cdot P\{\xi_{i_2} < x_{i_2}\} \cdots P\{\xi_{i_r} < x_{i_r}\}$$

其中x_{i_1}，x_{i_2}，x_{i_r}为任意实数，则称 ξ_1，ξ_2，…，ξ_n是**相互独立的**。

特别地，对二元随机变量(ξ, η)而言，当ξ，η相互独立时，即

$$P\{\xi > x, \eta < y\} = P\{\xi < x\} P\{\eta < y\}$$

对一切实数x，y皆成立，即有

$$F(x, y) = F_1(x) F_2(y).$$

若ξ，η是连续型随机变量，则其分布密度满足

$$f(x, y) = f_1(x) \cdot f_2(y)。$$

这里$F_1(x)$、$f_1(x)$和$F_2(x)$、$f_2(x)$分别是 ξ 和 η 的分布函数、分布密度。

设(ξ, η)是任意的一个二元随机变量，则

$$F_1(x) = P\{\xi < x\} = P\{\xi < x, \eta < +\infty\}$$
$$= F(x, \infty),$$
$$F_2(y) = P\{\eta < y\} = P\{\xi < +\infty, \eta < y\}$$
$$= F(\infty, y).$$

称$F_1(x)$及$F_2(y)$为二元随机变量(ξ,η)的**边际分布函数**。特别当(ξ,η)为连续时，

$$f_1(x)=\frac{d}{dx}F_1(x)=\frac{d}{dx}F(x,\infty)=\int_{-\infty}^{\infty}f(x,y)dy,$$

$$f_2(y)=\frac{d}{dx}F_2(y)=\frac{d}{dx}F(\infty,y)=\int_{-\infty}^{\infty}f(x,y)dx$$

它们分别是ξ和η的**边际分布密度**。

现在，我们转到讨论和条件概率有关的条件分布函数和条件分布密度问题。

设η为一随机变量，A为一事件，其概率$P(A)>0$，则称

$$\Phi(y|A)=P\{\eta<y|A\}$$

为η在条件A下的**条件分布函数**。若存在非负函数$f(y|A)$，使

$$\Phi(y|A)=\int_{-\infty}^{y}f(s|A)ds$$

则称$f(y|A)$为η在条件A之下的**条件分布密度**。

试虑二元随机变量(ξ,η)，$F(x,y)$为其分布函数。令$A_{\alpha\beta}=\{x-\alpha\leqslant\xi<x+\beta\}\quad(\alpha,\beta>0)$，而$B=\{\eta<y\}$，根据条件概率的定义

$$\begin{aligned}P\{BA_{\alpha\beta}\}&=P\{A_{\alpha\beta}\}P\{B|A_{\alpha\beta}\}\\&=(F(x+\beta,\infty)-F(x-\alpha,\infty))P\{B|A_{\alpha\beta}\}。\end{aligned}$$

因而

$$\begin{aligned}P\{B|A_{\alpha\beta}\}&=\frac{P\{BA_{\alpha\beta}\}}{F(x+\beta,\infty)-F(x-\alpha,\infty)}\\&=\frac{P\{x-\alpha\leqslant\xi<x+\beta,\eta<\eta\}}{F(x+\beta,\infty)-F(x-\alpha,\infty)}\end{aligned}$$

或者

$$\Phi(y|x-\alpha\leqslant\xi<x+\beta)=P\{\eta<y|x-\alpha\leqslant\xi<x+\beta\}$$

$$=\frac{P\{x-\alpha\leqslant\xi<x+\beta,\ \eta<y\}}{F(x+\beta,\infty)-F(x-\alpha,\infty)}$$

当$\alpha,\beta\to0^+$时，若上式极限存在，我们将极限值记作$\Phi(y|x)$，即

$$\Phi(y|x)=\lim_{\alpha,\beta\to0^+}\frac{F(x+\beta,y)-F(x-\alpha,y)}{F(x+\beta,\infty)-F(x-\alpha,\infty)}\text{.}$$

若存在非负函数$f(y|x)$，使下式成立

$$\Phi(y|x)=\int_{-\infty}^{y}f(s|x)ds$$

则称$\Phi(y|x)$和$f(y|x)$为在$\xi=x$条件下，η的**条件分布函数**和**条件分布密度。**

若设(ξ,η)是连续随机变量，其分布密度为$f(x,y)$，则

$$\Phi(y|x)=\lim_{\alpha,\beta\to0^+}\frac{\int_{x-\alpha}^{x+\beta}\int_{-\infty}^{y}f(s,t)dsdt}{\int_{x-\alpha}^{x+\beta}\int_{-\infty}^{\infty}f(s,t)dsdt}$$

$$=\frac{\int_{-\infty}^{y}f(x,t)dt}{\int_{-\infty}^{\infty}f(x,t)dt}=\frac{1}{f_1(x)}\int_{-\infty}^{y}f(x,t)dt\text{.}$$

所以有 $\quad f(y|x)=\dfrac{f(x,y)}{f_1(x)}$，

同理 $\quad f(x|y)=\dfrac{f(x,y)}{f_2(y)}$，

即 $\quad f(x,y)=f_1(x)f(y|x)=f_2(y)f(x|y)$

条件分布密度公式

边际分布密度函数和边际分布函数可表为

$$f_1(x)=\int_{-\infty}^{\infty}f_2(y)f(x|y)dy,$$

$$f_2(y)=\int_{-\infty}^{\infty}f_1(x)f(y|x)dx,$$

$$F_1(x)=\int_{-\infty}^{x}f_1(s)ds=\int_{-\infty}^{x}ds\int_{-\infty}^{\infty}f_2(y)f(s|y)dy$$

$$=\int_{-\infty}^{\infty}f_2(y)\Phi(x|y)dy,$$

$$f_2(y)=\int_{-\infty}^{y}f_2(s)ds=\int_{-\infty}^{y}ds\int_{-\infty}^{\infty}f_1(x)f(s|x)dx$$

$$=\int_{-\infty}^{\infty}f_1(x)\Phi(y|x)dx,$$

故得

$$\left.\begin{aligned}F_1(x)&=\int_{-\infty}^{\infty}\Phi(x|y)dF_2(y)\\F_2(y)&=\int_{-\infty}^{\infty}\Phi(y|x)dF_1(x)\end{aligned}\right|$$

全概率公式

$$f(x|y)=\frac{f(x,y)}{f_2(y)}=\frac{f_1(x)f(y|x)}{\int_{-\infty}^{\infty}f(x,y)ds}$$

$$=\frac{f_1(x)f(y|x)}{\int_{-\infty}^{\infty}f_1(x)f(y|x)dx}$$

同理可得

$$f(y|x)=\frac{f_2(y)f(x|y)}{\int_{-\infty}^{\infty}f_2(y)f(x|y)dy}$$

自然可将此公式称为贝叶斯公式。

全概率公式和贝叶斯公式是蒙特卡罗方法中常用的基本概率公式。

§6　随机变量的数字特征

在前面各节中我们介绍了有关随机现象的基本知识和工具。在本节，我们将讨论处理随机事件的一些基本统计方法和基本知识。有了这两个方面的准备，我们就可以进一步去掌握蒙特卡罗方法。

一、数学期望

某车间的机床工作方式是时而停止时而工作的，为了精确估计车间的电力负荷，需要知道同时工作的机床台数。为此，作了很多次观察。设总观察次数为N，发现有K台机床

工作的观察次数为m_K，$K=0$，1，2，…，M。这里 M 为总的机床台数。那么该车间中工作着的机床平均台数$\bar{n}$是

$$\bar{n}=\sum_{i=0}^{M}\frac{im_i}{N}=\sum_{i=0}^{M}i\omega_i$$

这里$\omega_i=\frac{m_i}{N}$是发现 i 台机床工作着的频率。但频率是随机地变动的，为了精确反映这个平均值，自然要用有 i 台机床工作的概率 p_i 来代替频率ω_i。从而我们得到随机变量ξ（同时工作的机床台数）的以概率为权的权平均值，这就是ξ的**数学期望**，记作$M\xi$，即

$$M\xi=\sum_{i=0}^{M}ip_i\ .$$

一般说来，设x_1，x_2，…，x_n，…是离散型随机变量ξ的可能取得值，而p_1，p_2，…，p_n，…分别表示ξ取这些值的概率，如果级数

$$\sum_{i=1}^{\infty}x_ip_i$$

绝对收敛，则将此级数的和值称为**离散型随机变量ξ的数学期望**，记为$M\xi$。

一般地，如果随机变量ξ的分布函数为$F(x)$，且积分

$$\int_{-\infty}^{\infty}xdF(x)\qquad\text{或}\qquad\int_{-\infty}^{\infty}xf(x)dx$$

绝对收敛，那么将其积分值称为**随机变量ξ的数学期望。**

例 1　设电话总机在长为 t 的时间内接到K次呼唤的概率为

$$P_K(t)=e^{-\lambda t}\frac{(\lambda t)^K}{K!}$$

求在长为 t 的时间内电话总机平均收到的呼唤次数。

解：平均次数等于下面的数学期望

$$M\xi=\sum_{K=0}^{\infty}KP_K(t)=\sum_{K=0}^{\infty}K\frac{(\lambda t)^K}{K!}\cdot e^{-\lambda t}$$

$$=e^{-\lambda t}\sum_{K=1}^{\infty}\frac{(\lambda t)^K}{(K-1)!}=e^{-\lambda t}\ \lambda t\cdot$$

$$\sum_{K=0}^{\infty}\frac{(\lambda t)^K}{K!}=\lambda t$$

例 2　设ξ在（a，b）上均匀分布，则

$$M\xi=\int_{-\infty}^{\infty}xdF(x)=\int_a^b\frac{x}{b-a}dx=\frac{a+b}{2}$$

例 3　对某一长度为 a 的物体进行测量，测量结果服从正态分布律，其分布密度为

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

所以，测量结果的平均值是：

$$M\xi=\int_{-\infty}^{\infty}xf(x)dx=a$$

例 4　从物理学中知道分子速度的绝对值的分布密度服从马克斯威尔分布律：

$$f(x)=\begin{cases}\dfrac{4x^2}{\alpha^3\sqrt{\pi}}\,e^{-\frac{x^2}{\alpha^2}} & (x>0)\\ 0 & (x\leqslant 0)\end{cases}$$

其中$\alpha>0$是常数。其分子的平均速度$\bar{v}$是

$$\bar{v}=\int_0^{\infty}x\frac{4x^2}{\alpha^3\sqrt{\pi}}e^{-\frac{x^2}{\alpha^2}}dx=\frac{2\alpha}{\sqrt{\pi}}.$$

由于我们定义了条件概率和条件分布函数（条件分布密度），因此我们可以类似地定义**条件数学期望**。

设$F(x|B)$是随机变量ξ对事件B的条件分布函数，则积分

$$M(\xi|B)=\int xdF(x|B)\ (\text{或}=\int xf(x|B)dx)$$

称为随机变量ξ对事件B的**条件数学期望**（$f(x|B)$是条件分布密度）。

设B_1，B_2，…，B_n是互不相容的事件，并且$\bigcup\limits_{K=1}^{n}B_K=U$是必然事件；而$F(x|B_1)$，$F(x|B_2)$，…，$F(x|B_n)$是随机变量$\xi$对这些事件相应的条件分布函数。那么，因为

$$\{\xi<x\}=\{\xi<x\}\cap[\bigcup_{K=1}^{n}B_K]$$

$$=\bigcup_{K=1}^{n}[\{\xi<x\}\cap B_K]$$

$$F(x)=P\{\xi<x\}=\sum_{K=1}^{n}P[\{\xi<x\}\cap B_K],$$

又因为

$$P[\{\xi<x\}\cap B_K]=P(B_K)P[\{\xi<x\}\mid B_K]$$
$$=P(B_K)F(x\mid B_K)$$

所以 $$F(x)=\sum_{K=1}^{n}P(B_K)F(x\mid B_K).$$

进而

$$M\xi=\int xdF(x)=\sum_{K=1}^{n}P(B_K)\int xdF(x\mid B_K)$$

$$=\sum_{K=1}^{n}P(B_K)M(\xi\mid B_K)=M\{M(\xi\mid B_K)\}\ .$$

这个公式在许多场合可以使数学期望的计算大为简化。

例 5　一个工人管 n 台同类型的机床，这些机床按等间距 a 布列在一条直线上。试求诸机床间的平均路经距离。

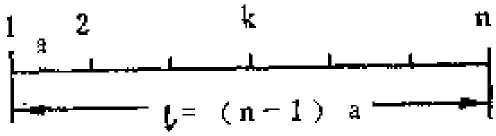

图1—6，a

把机床按顺序编号为 1，2，…n（如图 1—6，a）设每台机床需要照顾的概率均为$1/n$。B_K 表示“工人在第 K 台机床处”这一事件，这时第 i 台机床要求照顾（$1\leqslant i\leqslant n$），其路经距离为

$$\lambda_i^{(K)}=\begin{cases}(K-i)a & K\geqslant i\\(i-K)a & K<i\end{cases}$$

于是在B_K条件下的平均路经为

$$M(\lambda\mid B_K)=\frac{1}{n}\left[\sum_{i=1}^{K}(K-i)a+\sum_{i=K+1}^{n}(i-K)a\right]$$

$$=\frac{a}{2n}[2K^2-2(n+1)K+n(n+1)].$$

另一方面，$P(B_K)$也应为$1/n$，所以

$$M\lambda=M\{M(\lambda|B_K)\}=\frac{1}{n}\sum_{K=1}^{n}M(\lambda|B_K)$$

$$=\frac{a(n^2-1)}{3n}.$$

多元随机变量$\xi=(\xi_1, \xi_2, \cdots, \xi_n)$的数学期望定义为$M\xi(M\xi_1, M\xi_2, \cdots, M\xi_n)$，其中

$$M\xi_i=\overbrace{\int\cdots\int}^{n}x_i dF(x_1,x_2,\cdots,x_n)$$

$$=\int_{-\infty}^{\infty}x dF_i(x)$$

$F_i(x)=F(\infty,\infty,\cdots,\infty,\underset{i}{x},\infty,\cdots,\infty)$为边际分布函数.

二、方差

在研究随机变量时，数学期望表示随机变量的平均值，它是一个重要的数字特征。但是在许多实际问题中除了要知道随机变量的数学期望外，还需要进一步知道随机变量与其数学期望之间的偏差情况。例如，甲、乙两个射手记录了他们N次射击成绩。设d为击中点与靶心的距离（厘米），当$N=10$时，两个射手的射击记录如下表所示

	$d<1$	$1\leqslant d<2$	$2\leqslant d<3$	$3\leqslant d<4$	$4\leqslant d$
甲	6	3	1	0	0
乙	3	3	2	1	1

可以看出，尽管作为随机变量的甲，乙射手的击中点坐标其平均值（数学期望）都是靶心坐标，但是他们的“偏离程度”有着明显的不同。

我们用**方差**

$$D\xi=M(\xi-M\xi)^2=M\eta \qquad \eta=(\xi-M\xi)^2$$

来刻划随机变量ξ与其数学期望的偏离程度。若η的分布函数为$F_\eta(x)$，ξ的分布函数为$F_\xi(x)$，那么

$$D\xi=\int_{-\infty}^{\infty}x\,dF_\eta(x)$$

然而，通常我们并不知道$F_\eta(x)$，而只知道$F_\xi(x)$，为了解决求方差的计算问题，我们利用下面的定理

定理　如果$F_\xi(x)$是ξ的分布函数，$\varphi(x)$是连续函数，则

$$M\varphi(\xi)=\int_{-\infty}^{\infty}\varphi(x)\,dF_\xi(x)$$

证明（1）首先考虑$\varphi(x)$是x的严格增函数的情况

$$\eta=\varphi(\xi)$$

$$F_\eta(x)=P\{\eta<x\}=P\{\varphi(\xi)<x\}$$
$$=P\{\xi<\varphi^{-1}(x)\}=F_\xi[\varphi^{-1}(x)]$$

所以

$$M\varphi(\xi)=\int_{-\infty}^{\infty}x\,dF_\eta(x)=\int_{-\infty}^{\infty}x\,dF_\xi[\varphi^{-1}(x)]$$

$$= \int_{-\infty}^{\infty} \varphi(y) dF_\xi(y)$$

（2）对一般可微函数$\varphi(x)$有下列分解：

$$\varphi'(x) = \varphi_+'(x) - \varphi_-'(x)$$

所以
$$\varphi(x) = \varphi_+(x) - \varphi_-(x)$$

$$\varphi_\pm(x) = \int \varphi_\pm'(x) dx,$$

不妨设 $\varphi'_\pm(x) > 0$

则$\varphi_+(x)$，$\varphi_-(x)$为严格单调增函数。

$$M\varphi(\xi) = M(\varphi_+(\xi) - \varphi_-(\xi)) = M\varphi_+(\xi) - M\varphi_-(\xi)$$

$$= \int \varphi_+(x) dF_\xi(x) - \int \varphi_-(x) dF_\xi(x)$$

$$= \int [\varphi_+(x) - \varphi_-(x)] dF_\xi(x)$$

$$= \int \varphi(x) dF_\xi(x)$$

上面利用了数学期望的加法定理(见本节数学期望性质3)。

根据这个定理我们得到计算公式：

(1) $D\xi = M(\xi - M\xi)^2 = \int_{-\infty}^{\infty} (x - M\xi)^2 dF_\xi(x)$

(2) $D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 dF_\xi(x)$

$$= \int_{-\infty}^{\infty} [x^2 - 2xM\xi + (M\xi)^2] dF_\xi(x)$$

$$= \int_{-\infty}^{\infty} x^2 dF_\xi(x) - 2M\xi \int_{-\infty}^{\infty} x dF_\xi(x) + (M\xi)^2$$

$$\int_{-\infty}^{\infty} dF_\xi(x) = M\xi^2 - (M\xi)^2$$

(3) $M\xi^2 \geqslant (M\xi)^2$ 即 $\int_{-\infty}^{\infty} x^2 dF_\xi(x)$

$$\geqslant \left[\int_{-\infty}^{\infty} x dF_\xi(x) \right]^2$$

注意： 应该指出，不是所有的随机变量都存在有限的数学期望和方差的。例如分布函数为

$$F(x) = \begin{cases} 1 - \dfrac{1}{(1+x)^\alpha} & x \geqslant 0 \\ 0 & x < 0 \end{cases} \qquad (\alpha > 0)$$

的随机变量 ξ，其数学期望和方差如下：

$$M\xi = \int_{-\infty}^{\infty} x dF(x) = \int_0^{\infty} \frac{\alpha x}{(1+x)^{1+\alpha}} dx$$

$$D\xi = \alpha \int_0^{\infty} \frac{(x - M\xi)^2}{(1+x)^{1+\alpha}} dx$$

由微积分学知道，只有应 $\alpha > 1$ 时 $M\xi$ 才存在，而 $D\xi$ 只有当 $\alpha > 2$ 时才存在。

现在我们证明概率论中常用的重要不等式，即切贝雪夫（Чебышев）不等式。

切贝雪夫不等式 对任何具有有限方差的随机变量 ξ，都有

$$P\{|\xi - M\xi| \geqslant \varepsilon\} \leqslant \frac{D\xi}{\varepsilon^2} \qquad (\varepsilon > 0).$$

证明：设 $F(x)$ 是 ξ 的分布函数，则有

$$P\{|\xi-M\xi|\geqslant\varepsilon\}=\int_{|x-M\xi|\geqslant\varepsilon}dF(x)$$

$$\leqslant\int_{|x-M\xi|\geqslant\varepsilon}\frac{(x-M\xi)^2}{\varepsilon^2}dF(x)\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}(x-M\xi)^2dF(x)$$

$$=\frac{D\xi}{\varepsilon^2}.$$ 证毕.

该不等式说明方差愈小时，事件$\{|\xi-M\xi|\geqslant\varepsilon\}$的概率愈小。

推论 当$D\xi=0$时，ξ以概率1取常数$M\xi$，即$P\{\xi=M\xi\}=1$.

例1 试求均匀分布于区间(a, b)中的随机变量ξ的方差。

解：已知$M\xi=\frac{a+b}{2}$，$D\xi=M\xi^2-(M\xi)^2$，只要计算$M\xi^2$：

$$M\xi^2=\int_a^b x^2dF(x)=\int_a^b\frac{x^2}{b-a}dx$$

$$=\frac{1}{3(b-a)}(b^3-a^3)=\frac{b^2+ab+a^2}{3}$$

因此 $$D\xi=\frac{b^2+ab+a^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}.$$

例2 试求服从正态分布律

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

的随机变量ξ的方差。

解：因为$M\xi = a$

$$D\xi = \int_{-\infty}^{\infty}(x-a)^2 f(x)dx = \frac{1}{\sigma\sqrt{2\pi}}$$

$$\int_{-\infty}^{\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$

作变量代换 $z = \frac{x-a}{\sigma}$，可得

$$D\xi = \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz = \frac{\sigma^2}{\sqrt{2\pi}}$$

$$\left\{\left(-ze^{-\frac{z^2}{2}}\right)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz\right\} = \sigma^2$$

例 3　分子速度绝对値的分布密度由马克斯威尔分布律给出

$$f(x) = \begin{cases} \frac{4x^2}{a^3\sqrt{\pi}} e^{-\frac{x^2}{a^2}} & x>0 \\ 0 & x<0 \end{cases} \qquad (a>0)$$

求分子速度的方差、平均动能（假设分子的质量为m）及动能的方差。

解：分子速度的方差为

$$D\xi = M\xi^2 - (M\xi)^2 = \int_{-\infty}^{\infty} x^2 f(x)dx$$

$$- \left[\int_{-\infty}^{\infty} x f(x)dx \right]^2 = \frac{3}{2}\alpha^2 - \frac{4\alpha^2}{\pi}$$

$$= \left(\frac{3}{2} - \frac{4}{\pi}\right)\alpha^2$$

动能 $\eta = \frac{1}{2}m\xi^2$，于是平均动能为

$$M\eta = \int_{-\infty}^{\infty} \frac{m}{2} x^2 f(x)dx = \frac{3}{4} m\alpha^2$$

动能的方差为

$$D\eta = M\eta^2 - (M\eta)^2 = \int_{-\infty}^{\infty} \frac{m^2}{4} x^4 f(x)dx - (M\eta)^2$$

$$= \frac{3}{8} m^2 \alpha^4 .$$

三、数学期望和方差的性质

性质1 常数的数学期望和方差。

一个随机变量是常数C，意味着其分布函数为

$$F(x) = \begin{cases} 0 & x \leqslant c \\ 1 & x > c \end{cases}$$

(见图1—6，b)由此易推出 $M(c) = c$，$D(c) = 0$.

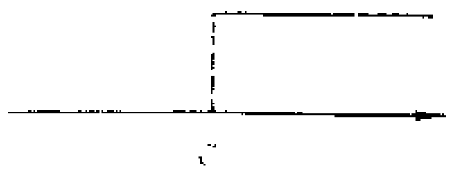

图1—6，b

性质2 设ξ为一随机变量，C为一常数，则 $M(C\xi) = CM\xi$，$D(C\xi) = C^2 D(\xi)$。

性质3 随机变量和的数学期望

设ξ_1和ξ_2是任意两个随机变量，则

$$M(\xi_1+\xi_2)=M\xi_1+M\xi_2 .$$

证明：设(ξ_1, ξ_2)，ξ_1及ξ_2的分布函数分别为$F(x,y)$，$F_1(x)$，$F_2(y)$。注意到

$$F_1(x)=P\{\xi_1<x\}=P\{\xi_1<x,\xi_2<\infty\}=F(x,\infty),$$

$$F_2(y)=P\{\xi_2<y\}=P\{\xi_1<\infty,\xi_2<y\}=F(\infty,y),$$

因而有

$$M(\xi_1+\xi_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)dF(x, y)$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xdF(x,y)+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}ydF(x,y)$$

$$=\int_{-\infty}^{\infty}xdF(x,\infty)+\int_{-\infty}^{\infty}ydF(\infty,y)$$

$$=\int_{-\infty}^{\infty}xdF_1(x)+\int_{-\infty}^{\infty}ydF_2(y)=M\xi_1+M\xi_2 .$$

性质4.1 当$x=M\xi$时，$M(\xi-x)^2$为最小。

证明 因为$M(\xi-x)^2=M\xi^2-2xM\xi+x^2=f(x)$，

$$\frac{df(x)}{dx}=2x-2M\xi \quad ,$$

当 $x=M\xi$ 时有 $\dfrac{df}{dx}=0$，$\dfrac{d^2f}{dx^2}=2>0$，性质得证。

性质4.2 当 $g(x)=M(\eta|\xi=x)$ 时，$M[\eta-g(x)]^2$ 为最小。

证明　因为

$$M[\eta-g(x)]^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[y-g(x)]^2 f(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}f_1(x)dx\int_{-\infty}^{\infty}[y-g(x)]^2 f(y|x)dy,$$

由性质4.1知积分

$$\int_{-\infty}^{\infty}[y-g(x)]^2 f(y|x)dy \text{ 在 } g(x)=M(\eta|\xi=x)\text{时为}$$

最小。

性质5　随机变量和的方差

设ξ,η为任意的两个随机变量，则

$$D(\xi+\eta)=D\xi+D\eta+2b_{12},$$

其中b_{12}是ξ与η的相关矩（其意义见证明）。

证明　$D(\xi+\eta)=M[\xi+\eta-M(\xi+\eta)]^2$

$$=M(\xi-M\xi+\eta-M\eta)^2=M(\xi-M\xi)^2$$

$$+2M[(\xi-M\xi)(\eta-M\eta)]+M(\eta-M\eta)^2$$

$$=D\xi+D\eta+2M[(\xi-M\xi)(\eta-M\eta)]$$

$$=D\xi+D\eta+2b_{12}$$

其中　$b_{12}=M[(\xi-M\xi)(\eta-M\eta)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}$

$$(x-M\xi)(y-M\eta)dF(x,y).$$

定义　设$b_{11}=D\xi$，$b_{22}=D\eta$，则

$$r_{12}=\frac{b_{12}}{\sqrt{b_{11}b_{22}}}$$

称为随机变量ξ，η的**相关系数**。

注 1：$|r_{12}|\leqslant 1$.

事实上，根据著名的哥西（Cauchy）不等式知

$$\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-M\xi)(y-M\eta)dF(x,\ y)\right|^2$$

$$\leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-M\xi)^2dF\cdot\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-M\eta)^2dF.$$

注 2：若ξ,η相互独立，则$r_{12}=0$，并且有

$$D(\xi+\eta)=D\xi+D\eta.$$

事实上，此时有$F(x,y)=F_1(x)\cdot F_2(y)$

$$b_{12}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-M\xi)(y-M\eta)dF_1(x)F_2(y)$$

$$=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}(x-M\xi)dF_1(x)\right]\cdot$$

$$\cdot(y-M\eta)dF_2(y)=0.$$

§7 极 限 定 理

在本节，我们要研究大量试验所服从的宏观规律性。若将随机事件的个别试验视为一种微观现象，如个别分子的速度，一次试验的结果等等，但是大量的试验结果在很多情况下有着确定的统计规律，如容器中气体的压力和温度等等。

一、大数定律

切贝雪夫大数定理 设ξ_1，ξ_2，…，ξ_n，…是由两两相

互独立的随机变量所构成的序列，每一随机变量都有有限的方差，并且它们被同一常数所界：

$$D\xi_1 \leqslant C, \quad D\xi_2 \leqslant C, \quad \cdots, \quad D\xi_n \leqslant C, \quad \cdots$$

则对任意的$\varepsilon>0$，皆有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{K=1}^{n}\xi_K - \frac{1}{n}\sum_{K=1}^{n}M\xi_K\right| < \varepsilon\right\} = 1.$$

证明：因为$\{\xi_K\}$是两两独立的，故

$$D\left(\frac{1}{n}\xi_K\right) = \frac{1}{n^2}\sum_{K=1}^{n}D\xi_K \leqslant \frac{C}{n}$$

再由切贝雪夫不等式得

$$P\left\{\left|\frac{1}{n}\sum_{K=1}^{n}\xi_K - \frac{1}{n}\sum_{K=1}^{n}M\xi_K\right| < \varepsilon\right\}$$

$$= 1 - P\left\{\left|\frac{1}{n}\sum_{K=1}^{n}\xi_K - \frac{1}{n}\sum_{K=1}^{n}M\xi_K\right| \geqslant \varepsilon\right\}$$

$$\geqslant 1 - \frac{1}{\varepsilon^2}D\left(\frac{1}{n}\sum_{K=1}^{n}\right)\xi_K) \geqslant 1 - \frac{C}{n\varepsilon^2} \longrightarrow 1 \quad (n\to\infty)$$

由此定理得证。

贝努里定理 设μ是n次独立试验中事件A出现的次数，而p是事件A出现的概率，则对任意的$\varepsilon>0$，都有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1 .$$

证明：令ξ_K代表第K次试验的结果

$$\xi_K = \begin{cases} 1 & A\text{ 出现} \\ 0 & A\text{ 不出现} \end{cases}$$

则$M\xi_K = p$，$D\xi_K = p(1-p)$，于是本定理成为上面切贝雪夫

定理的特例。

辛钦(Хинчин)定理 设ξ_1，ξ_2，…，ξ_n，…均为相互独立相同分布的随机变量，且具有有限的数学期望($a=M\xi_n$，$n=1,2,\cdots$)，则对任意的$\varepsilon>0$，有

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{K=1}^{n}\xi_K-a|<\varepsilon\}=1.$$

证明略。当ξ_K有有限方差时，本定理又是切贝雪夫大数定理的特例。

以上三个定理都引导出一个随机变量序列的收敛性概念。

定义 如果随机变量序列$\{\xi_n\}$和随机变量ξ，对任意的$\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|\xi_n-\xi|\geqslant\varepsilon\}=0$$

则称**序列$\{\xi_n\}$以概率收敛于ξ**。

我们还可引进更强的收敛性概念

定义 如果随机变量序列$\{\xi_n\}$和随机变量ξ,满足

$$P\{\xi_n\xrightarrow[n\to\infty]{}\xi\}=1$$

则称**序列$\{\xi_n\}$以概率1收敛于ξ**。

若序列$\{\xi_n\}$以概率收敛于ξ，也可以写成

$$\lim_{n\to\infty}P\{|\xi_n-\xi|<\varepsilon\}=1.$$

换言之，对任何正数$\eta>0$，存在充分大的N，当$n>N$时有

$$P\{|\xi_n-\xi|<\varepsilon\}>1-\eta.$$

若序列$\{\xi_n\}$以概率1收敛于ξ，可以叙述为,对任意正数$\varepsilon>0$，存在充分大的N，当$n>N$时有

$$P\{|\xi_n-\xi|<\varepsilon\}=1.$$

这里,我们不加证明地给出两个加强的结果。

柯尔莫哥洛夫(Колмогоров)定理(加强的大数定理) 设ξ_1, ξ_2, …, ξ_n, …是一个相互独立的随机变量序列，且$\sum_{n=1}^{\infty}\frac{D\xi_n}{n^2}<\infty$，则有

$$P\{\frac{1}{n}\sum_{K=1}^{n}\xi_K-\frac{1}{n}\sum_{K=1}^{n}M\xi_K\xrightarrow[n\to\infty]{}0\}=1.$$

定理 设ξ_1, ξ_2, …, ξ_n, …是一个相互独立具有相同分布的随机变量序列，则

$$P\{\frac{1}{n}\sum_{K=1}^{n}\xi_K-\frac{1}{n}\sum_{K=1}^{n}M\xi_K\xrightarrow[n\to\infty]{}0\}=1$$

的充要条件是随机变量序列$\{\xi_n\}$的数学期望存在。

二、中心极限定理

现在，我们给出几个有用的中心极限定理，它们的证明要涉及较多的数学工具，因而不在这里介绍。对独立的随机变量序列$\{\xi_n\}$，引进记号：

$$a_K=M\xi_K,\quad b_K^2=D\xi_K,\quad K=1,\ 2,\ \cdots$$

$$B_n^2=\sum_{K=1}^{n}b_K^2$$

定理 如果相互独立的随机变量序列ξ_1, ξ_2, …, ξ_n, …，对任何$\tau>0$满足林德伯尔格(Линдеберг)条件：

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{K=1}^{n}\int_{|x-a_K|\geqslant\tau B_n}(x-a_K)^2dF_K(x)=0,$$

则当$n\to\infty$时，对一切x，有

$$P\left\{\frac{1}{B_n}\sum_{K=1}^{n}(\xi_K-a_K)<x\right\}\longrightarrow\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{z^2}{2}}dz.$$

定理 如果相互独立的随机变量ξ_1，ξ_2，…，ξ_n，…具有相同分布，并且有有限的非零方差，则当$n\to\infty$时，对一切x，有

$$P\left\{\frac{1}{B_n}\sum_{K=1}^{n}(\xi_K-M\xi_K)<x\right\}\longrightarrow\frac{1}{\sqrt{\pi 2}}\int_{-\infty}^{x}e^{-\frac{z^2}{2}}dz,$$

这里 $B_n=\sqrt{n}\,b$，$M\xi_K=a$，$b^2=D\xi_K$，$K=1$，2，…。

例 1 对敌人防御地段进行100次的射击，至每次射击中，炮弹命中数的数学期望为4，命中数的均方差为1.5。求当射击100次时，有380颗到420颗炮弹命中目标的概率的近似值。

解： $a=4$，$b=1.5$，$B_n^2=100\times(1.5)^2=225$，

令 $$\xi=\sum_{K=1}^{100}\xi_K$$

于是 $$P\{380\leqslant\xi<420\}\approx P\left\{-\frac{20}{15}\leqslant\frac{\xi-400}{\sqrt{225}}<\frac{20}{15}\right\}$$

$$\approx\int_{-\frac{25}{15}}^{\frac{20}{15}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\approx 0.8164.$$

例 2 对长度为a的物体进行测量，误差在±0.5范围内。假设测量所得的值服从均匀分布，即$M\xi=a$，$D\xi=\frac{1}{12}=$

b^2。进行1200次独立测量，取测量的平均值作为结果。问结果精度不超过0.01的概率。

解法同上。答案：0.7698。

第二章　蒙特卡罗方法概述

本章首先讨论蒙特卡罗方法的基本思想，扼要介绍该方法的主要概念，主要问题。其次，将着重讨论用数学方法如何产生伪随机数的问题。

§1　蒙特卡罗方法的基本思想和一般过程

一、首先我们通过以下三个例子说明该方法的基本思想。

例 1　工厂为了确定某产品的合格率，常常抽查其中一部分（N个）产品，逐个进行检验。如果有 n 个合格品，那末自然认为合格率 p 为

$$p \approx n/N,$$

人们从经验中知道，N 数目越大，则 n/N 作为合格率的估计值就越准确。

例 2　蒲丰氏问题

在十九世纪后期，曾有很多人以任意投掷一根针到地面上，将针与地面上两条平行线相交的频率作为针与该两条平行线相交概率的近似值，然后根据这一概率的理论值

$$p = \frac{2l}{\pi a},$$

求出圆周率 π 的近似值：

$$\pi = \frac{2l}{a} \cdot \frac{1}{p} \approx \frac{2l}{a}\left(\frac{N}{n}\right),$$

这里，$2a$ 为平行线之间的距离，$2l(l<a)$是针的长度，N 是投针次数，n 是相交次数。这就是著名的蒲丰氏问题。一些人进行了实验，其结果列于下表：

实　验　者	年　份	投针次数	π的实验值
沃尔弗（Wolf）	1850	5000	3.1596
斯密思（Smith）	1855	3204	3.1553
福克斯（Fox）	1894	1120	3.1419
拉查里尼（Lazzarini）	1901	3408	3.1415929

例 3　射击问题

设 r 表示射击运动员的弹着点到靶心的距离，$g(r)$ 表示得分数（环数），分布密度函数 $f(r)$表示该运动员的弹着点分布，它反映运动员的射击水平。积分

$$\langle g\rangle = \int_0^{\infty} g(r)f(r)dr$$

表示这个运动员的射击成绩。用概率论语言说，$\langle g\rangle$ 就是随机变量 $g(r)$的数学期望。现在，让该运动员射击 N次，弹着点依次为$\{r_1, r_2, \cdots, r_N\}$，则自然地认为$N$次射击得分的平均值为

$$\bar{g}_N = \frac{1}{N}\sum_{n=1}^{N} g(r_n),$$

它近似地代表了运动员的成绩。换言之，g_N是上述积分的一个估计值。

以上三个例子中，一个是求合格率，一个是求数π，一个是求积分。它们的具体内容各不相同，但都是通过实际"试验"的方法，得到某种事件出现的频率，再进行统计平均以求得其近似值。廿世纪四十年代，由于电子计算机的出现，人们有可能对许多实际试验用计算机来模拟，这种模拟称为数字模拟。由于计算机具有高速度和大容量的特点，使数字模拟可以代替许多实际上非常庞大而复杂的试验，而在计算机上迅速地完成，同时将所得的结果立即进行统计处理。所以蒙特卡罗方法的最独特的性质就是用数学方法在电子计算机上实现数字模拟试验。

二、用蒙特卡罗方法解题的一般过程。其主要步骤是：

（1） 构造或描述问题的概率过程；

（2） 实现从已知概率分布的抽样；

（3） 建立各种统计量的估计。

下面我们详述这三个步骤。

（1） 对于本身就具有随机性质的问题，如粒子输运问题，主要是正确地描述和模拟这个概率过程。对于本来不是随机性质的确定性问题，比如计算定积分、解线性方程组及偏微分方程边值问题等，要用蒙特卡罗方法求解，就必须事先构造一个人为的概率过程，它的某些参量正好是所要求的问题的解。

（2） 有了明确的概率过程后，为了实现过程的数字模拟，必须实现从已知概率分布的随机数的抽样。以射击问题为例，要得到积分$\langle g\rangle$的估计值g_N，关键在于得到弹着点

r 的序列：$r_1, r_2, \cdots, r_N$，从而得到 $g(r_1), g(r_2), \cdots, g(r_N)$。但 r 的产生是由运动员的弹着点分布密度函数 $f(r)$ 决定的，这是从运动员本身射击经验历史总结的规律，因此上述 r 的序列必须遵从分布律 $f(r)$。这种根据已知分布律（密度或分布密度）产生的随机变量 r 的具体值 $r_1, r_2, \cdots, r_N$，称为分布律 $f(r)$ 的一个**子样**，r_k 为子样的一个**元素**。这就是从已知分布律 $f(r)$ 实现抽样的问题。

最简单、最重要、最基本的一个概率分布是（0，1）上的均匀分布（或称矩形分布）。**随机数**就是具有这种均匀分布的随机变量。许多其他复杂的分布可以用数学方法由它产生。在计算机上可以用物理方法直接产生随机数，但是价格昂贵，不能重复，使用不便。另一种方法是用数学方法按一定的递推关系产生子样，它与真正的随机数序列不同，称为**伪随机数**（*Pseudo-random Numbers*）。伪随机数与真正的随机数具有相近的性质，可把它作为真正的随机数来使用。

（3） 一般说来，构造了概率模型并能从中抽样后，即能实现数字模拟试验后，我们就要确定一个随机变量，作为所要求的问题的解的某种数字量的估计量。如果这个随机变量的数学期望正好是所求问题的解，我们称这种估计量为**无偏估计**。

最后，作为一个例题，我们将用蒙特卡罗方法解蒲丰氏问题。首先，投针问题本身就是具有随机性质的问题，因此现在的任务是正确地描述与模拟这个问题。

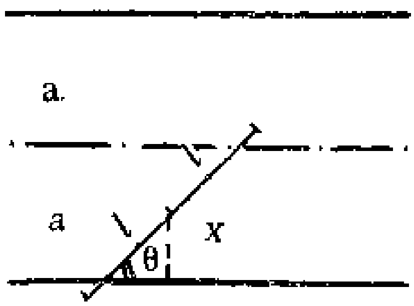

图2—1，a

事实上，正如图 2 — 1 a，

所示，平面上一根针的位置可以用针中心的坐标 x 和针与平行线的夹角 θ 来决定。任意投针，意味着 x 与 θ 都是任意取的，只是 θ 的范围可限于 $[0,\pi]$，而 x 的范围可限于 $[0,a]$。此时相交的条件为：

$$x \leqslant l\sin\theta，\quad 0 \leqslant x \leqslant a，\quad 0 \leqslant \theta \leqslant \pi。$$

由随机变量 $\{x,\theta\}$ 在各自取值范围内取值的随意性，可知 x 和 θ 各自的分布密度应为：

$$f_1(x)=\begin{cases}\dfrac{1}{a} & 0 \leqslant x \leqslant a\\ 0 & 其他\end{cases}，$$

$$f_2(\theta)=\begin{cases}\dfrac{1}{\pi} & 0 \leqslant \theta \leqslant \pi\\ 0 & 其他\end{cases}。$$

而 $\{x,\theta\}$ 的分布密度为 $f(x,\theta)=f_1(x)f_2(\theta)$。实现从分布 $f(x,\theta)$ 中抽样的问题可以转化为

$$x=a\xi_1，\quad \theta=\pi\xi_2$$

这里，ξ_1 和 ξ_2 均为（0，1）上均匀分布的独立的随机数。在电子计算机上 ξ_1 和 ξ_2 可分别用两个伪随机数来代替。

这样每次投针试验，实际上变成从两个均匀分布的随机变量中抽样取得 (x_i,θ_i)，然后定义随机变量 $s(x,\theta)$ 如下：

$$s(x_i,\theta_i)=\begin{cases}1 & 当\ x_i \leqslant l\sin\theta_i\\ 0 & 其他\end{cases}$$

如果投针 N 次，那么

$$\bar{s}_N=\frac{1}{N}\sum_{i=1}^{N}s(x_i,\theta_i)$$

是相交概率 p 的估计值，事实上

$$p=\iint s(x,\theta)f(x,\theta)dxd\theta=\int_0^{\pi}\frac{d\theta}{\pi}\cdot\int_0^{l\sin\theta}\frac{1}{a}dx$$

$$=\frac{2l}{a\pi}$$

于是有

$$\pi=\frac{2l}{ap}\approx\frac{2l}{a\bar{s}_N}$$

在计算机上进行计算所得结果列于下表：

a,l / π / N	$a=3$, $l=2.5$	$a=4$, $l=3$	$a=5$, $l=4$	$a=10$, $l=9.99$
20,000	3.1152	3.1348	3.1305	3.1308
100,000	3.1440	3.1482	3.1488	3.1467
200,000	3.1478	3.1480	3.1485	3.1421
500,000	3.1432	3.1436	3.1436	3.1435

现在我们进一步用比较抽象的概率语言，把蒙特卡罗方法解题的过程归纳如下：构造一个概率空间 $(\Omega,\mathscr{F},P)$，其中 Ω 是一个事件集合，$\mathscr{F}$ 是集合 Ω 的子集合构成的集合（满足第一章§2中所列的条件），P 是在 $\mathscr{F}$ 上建立的概率；在这个概率空间中，选取一个随机变量 $\theta(\omega)$，$\omega\in\Omega$，使其数学期望

$$\Theta = \int_{\Omega} \theta(\omega) P(d\omega)$$

正好是所要求的解 Θ，然后取 $\theta(\omega)$ 的子样的算术平均值为 Θ 的近似值。

实验题 1：试在计算机上作蒲丰氏问题的模拟（提示：利用 *Basic* 或 *Fortran* 内部随机数 *RAN*(I)作为随机数）。

§2　蒙特卡罗方法的收敛性和误差估计

一、收敛性及收敛速度

上面已提到，蒙特卡罗方法常常以随机变量 $\theta(\omega)$ 的子样 $\theta(\omega_1),\theta(\omega_2),\cdots,\theta(\omega_N)$ 的算术平均值

$$\bar{\theta}_N = \frac{1}{N}\sum_{n=1}^{N}\theta(\omega_n)$$

作为求解 Θ 的近似值。

由大数定律或辛钦定理知道，相互独立、相同分布、具有有限数学期望的随机变量序列$\{\theta(\omega_n),\ n=1,2,\cdots,N\}$，对任何 $\varepsilon>0$，有

$$\lim_{N\to\infty} P\{|\bar{\theta}_N-\Theta|<\varepsilon\}=1。$$

由加强的大数定律，则有

$$P\{\lim_{N\to\infty}\bar{\theta}_N=\Theta\}=1.$$

由此推得，在蒙特卡罗方法中，随机变量 $\theta(\omega)$ 的子样的算术平均值 $\bar{\theta}_N$，当 $N\to\infty$时，以概率 1 收敛到 Θ。

按照中心极限定理，若要求 $\theta(\omega)$有有限方差 $\sigma\neq 0$，即

$$0\neq\sigma^2=\int_\Omega(\theta(\omega)-\Theta)^2P(d\omega)<+\infty,$$

那么，当 $N\to\infty$时，随机变量

$$Y_N=(\bar{\theta}_N-\Theta)\sqrt{N}/\sigma,\quad MY_N=0,\quad DY_N=1.$$

渐近于正态分布，即有

$$P\{Y_N<x_\alpha\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_\alpha}e^{-\frac{x^2}{2}}dx,$$

$$P\{|Y_N|<x_\alpha\}=P\{Y_N<x_\alpha\}-P\{Y_N<-x_\alpha\}$$

$$=P\left\{|\bar{\theta}_N-\Theta|<\frac{x_\alpha\sigma}{\sqrt{N}}\right\}$$

$$\approx\frac{2}{\sqrt{2\pi}}\int_0^{x_\alpha}e^{-\frac{x^2}{2}}dx=1-\alpha$$

这表明，不等式

$$|\bar{\theta}_N-\Theta|<\frac{x_\alpha\sigma}{\sqrt{N}}$$

大约以概率 $1-\alpha$ 成立。通常 α 称为**置信度**，$1-\alpha$ 称为**置信水平**。α 与 x_α 一一对应。上式表明，$\bar{\theta}_N$ 收敛到 Θ 的速度的阶为 $O(N^{-\frac{1}{2}})$。常用的几组 α 和 x_α 如下：

α	x_α
0.5	0.6745
0.05	1.96
0.01	3

二、误差估计

如果 $\sigma \neq 0$，则从前面的讨论得知，蒙特卡罗方法的误差 ε 为

$$\varepsilon = x_\alpha \sigma / \sqrt{N}$$

并且误差估计

$$|\bar{\theta}_N - \Theta| < \varepsilon = x_\alpha \sigma / \sqrt{N}$$

约以概率 $1-\alpha$ 成立。所以置信度 α 确定以后，误差完全由 σ 和 $\sqrt{N}$ 决定。要想减小 ε，只有增大 N，或者减小方差 σ。若固定 σ，则要想提高精度一位数字（即10倍），就要将 N 提高100倍。因此单纯增大 N，不一定是个好办法。

减小误差 ε 的另一种途径是减小方差 σ。设某种方法将方差减小为 σ_1，那么在同一置信度要求下，N 可以减小至 N_1，即

$$\frac{\sigma}{\sqrt{N}} = \frac{\sigma_1}{\sqrt{N_1}}, \quad N_1 = N\frac{\sigma_1^2}{\sigma^2}。$$

但是我们不能单从这个关系决定选择那种方法，还应该考虑观察子样中一个元素的工作量（或者费用，机时）C。设新的方法相应的工作量为 C_1，于是总工作量 W 和 W_1 分别为 NC 和 N_1C_1，它们的比值应为

$$\frac{W_1}{W} = \frac{N_1C_1}{NC} = \frac{\sigma_1^2 C_1}{\sigma^2 C}$$

所以取 $\sigma^2 C$ 越小的方法，就越经济。

最后，我们指出将蒙特卡罗方法和解析（数值）方法组合的技巧。用这种方法可以有效地减小方差 σ，从而提高方法的精度。考虑计算定积分

$$\theta = \iint_{V_2} g(x,y) f(x,y) dx dy$$

其中 $f(x,y)$是 V_2 上联合密度函数，即有

$$f(x,y) \geq 0 \text{ , } \iint_{V_2} f(x,y) dx dy = 1$$

令

$$f_1(x) = \int f(x,y) dy$$

$$f(x,y) = f_1(x) f_2(y|x),$$

$f_1(x)$和 $f_2(y|x)$分别为边际分布密度和条件分布密度。记条件 $\xi = x$ 之下的数学期望及方差为

$$\theta_x = \theta(x) = \int g(x,y) f_2(y|x) dy$$

$$\sigma_x^2 = \int (g(x,y) - \theta_x)^2 f_2(y|x) dy_{\circ}$$

利用通常的蒙特卡罗方法计算 θ，是从分布 $f(x,y)$ 抽取子样序列 $(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)$后计算平均值

$$\bar{\theta}_N = \frac{1}{N} \sum_{k=1}^{N} g(x_k, y_k)$$

作为 θ 的估计值。这种方法的误差取决于方差：

$$\sigma^2 = \iint_{V_2} (g(x,y) - \theta)^2 f(x,y) dx dy_{\circ}$$

另一种方法是利用关系式

$$\theta=\int \theta_x f_1(x)dx,$$

如果我们能用解析（或数值）方法精确地计算 θ_x 的话，我们便可以从分布 $f_1(x)$ 中抽取子样序列 $x_1,x_2,\cdots,x_N$ 后，计算平均值

$$\bar{\theta}_{xN}=\frac{1}{N}\sum_{k=1}^{N}\theta(x_k)$$

作为 θ 的估计值。这种方法的误差只取决于方差

$$\sigma_1{}^2=\int(\theta_x-\theta)^2 f_1(x)dx$$

这两个方差满足下列关系：

$$\sigma^2=\iint(g(x,y)-\theta)^2 f(x,y)dxdy$$

$$=\iint(g(x,y)-\theta_x+\theta_x-\theta)^2 f_1(x)f_2(y|x)dxdy$$

$$=\iint(g(x,y)-\theta_x)^2 f_1(x)f_2(y|x)dydx$$

$$+\iint(\theta_x-\theta)^2 f_1(x)f_2(y|x)dydx$$

$$+2\iint(g(x,y)-\theta_x)(\theta_x-\theta)f_1(x)f_2(y|x)dydx$$

$$=\iint\sigma_x{}^2 f_1(x)dx+\int(\theta_x-\theta)^2 f_1(x)dx+0$$

$$= \int \sigma_x{}^2 f_1(x)dx + \sigma_1{}^2$$

由此可见$\sigma^2 \geq \sigma_1{}^2$.

从上面的分析中，我们得到一个一般的原则：我们只对问题中难于用解析（或数值）方法处理的部分，才利用蒙特卡罗方法计算，其他部分尽可能多地利用解析方法。这样可以减小方差和提高精度。

实验题 2　试计算一个定积分

$$\iint_{\Omega} g(x,y)dxdy$$

其中Ω为正方形(提示:利用内部函数 $RAN(I_1)$,$RAN(I_2)$).

§3　随机数和伪随机数

一、各种各样分布的随机变量中最简单、最基本的随机变量是矩形分布密度的随机变量。用适当的数学方法可以由它形成各种其他分布的随机变量。这是蒙特卡罗方法最基本、最经常使用的手段。因此，如何在电子计算机上产生这种矩形分布的随机变量便成了我们的基本问题，随之而来的是这样产生随机变量的质量问题。这些就是我们今后所关心的问题。

定义　从分布密度函数为

$$f(x) = \begin{cases} 1 & 0 \leqslant x \leqslant 1 \\ 0 & x \notin [0,1] \end{cases}$$

的随机变量总体中抽取的子样 $x_1, x_2, \cdots, x_N$ 称为**随机数序列**，其中每个个体称为**随机数**。

今后我们将用带下标的希腊字母序列表示随机数序列：$\xi_1, \xi_2, \cdots, \xi_N$。根据定义，我们可以断定随机数序列是相互独立的、具有相同矩形分布的随机变量序列。于是，随机数具有非常重要的性质：对任意 S 个随机数组成的序列，作为 S 维空间的点 $(\xi_1, \xi_2, \cdots, \xi_S)$，在 S 维空间的单位方体 G_S 上均匀分布，即设

$$G_S = \{(x_1, \cdots, x_S) \mid 0 \leqslant x_i \leqslant 1 \quad i = 1, 2, \cdots, S\}$$

则有

$$P\{\xi_i \leqslant x_i \ \ i = 1, 2, \cdots, S \mid 0 \leqslant x_i \leqslant 1 \ \ i = 1, 2, \cdots, S\}$$

$$= \prod_{i=1}^{S} x_i。$$

为了产生随机数，过去有人用各种方法将随机数制表。但是制表的容量有限，而且放到计算机里去要消耗大量内存，因而不适于大规模的试验。也可以用物理方法制造白色噪声发生器，然后数字化后得到随机数，但是价格昂贵，而且也不能重复，因此也不适用。

二、鉴于上述情况，人们致力于用数学方法来产生随机数，其基本思想是利用一种递推公式：

$$\xi_{n+k} = T(\xi_n, \xi_{n+1}, \cdots, \xi_{n+k-1}),$$

对于给定的初始值 $\xi_1, \xi_2, \cdots, \xi_k$，逐个地产生 ξ_{k+n}，$n = 1, 2, \cdots$。经常遇到的是 $k = 1$ 的情况，这时用一种单步递推公式：

$$\xi_{n+1} = T(\xi_n)$$

对于给定的初始值 ξ_1，逐个地产生 $\xi_2, \xi_3, \cdots$。

这种用数学方法产生的随机数存在两个问题：

1　整个随机数序列完全由T和初始值唯一确定，严格地说它不满足随机数的相互独立的要求。

2　既然随机数序列是用递推公式确定的，而在电子计算机上所能表示的[0,1]上的数又是有限多的，因此，这样任何递推公式不能不出现重复的循环。一旦这种情况产生，数列便出现了周期性的循环，这又同随机数的要求是相矛盾的。

由于这两个原因，常将用数学方法所产生的随机数称为**伪随机数**。

关于伪随机数的第一个问题，我们可以通过小心地选择T来使所产生的数列近似地满足独立性、均匀性要求；关于第二个问题则不是本质的，因为我们用蒙特卡罗方法解决任何问题时只进行有限次试验，只要其次数不超过伪随机数出现循环的长度就可以了。现在我们定量地讨论伪随机数的几个有关问题。

1　伪随机数的周期和最大容量

由于伪随机数序列一定会出现周期性循环现象，因此，在使用伪随机数时不可能无限制地继续下去，就是说由伪随机数所组成的子样容量应有一个最大限度。超过这个限度将出现伪随机数的重复。发生周期性循环现象的伪随机数个数，称为伪随机数的周期。设第一次出现重复的伪随机数的下标为 n' 和 n''（即 $\xi_{n'}=\xi_{n''}$ 第一次出现），则 $n''-n'$ 称为周期，n'' 称为最大容量。

2　伪随机数的均匀性

我们只考虑伪随机数序列 $\xi_1,\xi_2,\cdots,\xi_n$ 全体作为子样时的均匀性问题，其中 n 为伪随机数的最大容量。对于任意的 $x\in[0,1]$，令 $N_n(x)$ 表示该序列中适合不等式

$$\xi_i < x, \quad 1 \leqslant i \leqslant n$$

的个数，则

$$\delta(n) = \sup_{0 \le x \le 1} \left| \frac{N_n(x)}{n} - x \right|,$$

标志伪随机数序列 $\xi_1, \xi_2, \cdots, \xi_n$ 的均匀程度，简称为**均匀偏度**。很明显，对固定的 n 而言，$\delta(n)$的值越小均匀性越好。

将伪随机数序列 $\xi_1, \xi_2, \cdots, \xi_n$，按由小到大重新排列

$$\xi_1' \leqslant \xi_2' \leqslant \cdots \leqslant \xi_n',$$

并规定 $\xi_0' = 0$，$\xi_{n+1}' = 1$。则根据 $\delta(n)$的定义，我们有

$$\delta(n) = \sup_{0 \le x \le 1} \left| \frac{N_n(x)}{n} - x \right| = \sup_{\xi_i' < x \le \xi_{i+1}'} \left| \frac{i}{n} - x \right|$$

$$= \max_{0 \le i \le n} \left\{ \left| \xi_i' - \frac{i}{n} \right|, \left| \xi_{i+1}' - \frac{i}{n} \right| \right\}.$$

于是我们有下列性质：

① $\xi_1' - \xi_0' \leqslant \delta(n)$，

$\xi_{i+1}' - \xi_i' \leqslant 2\delta(n)$，$i = 1, 2, \cdots, n-1$，

$\xi_{n+1}' - \xi_n' \leqslant \delta(n)$。

关于第一、三不等式是显然的，至于第二个不等式只要注意

$$\left| \xi_{i+1}' - \xi_i' \right| \leqslant \left| \xi_{i+1}' - \frac{i}{n} \right| + \left| \xi_i' - \frac{i}{n} \right| \leqslant 2\delta(n).$$

② $\delta(n) \geq \frac{1}{2n}$

为证明此不等式只需将①中各不等式相加。

③关于最佳伪随机数分布

伪随机数序列 $\xi_1, \xi_2, \cdots, \xi_n$ 所对应的偏度 $\delta(n)$ 达到下界的充要条件是它取遍如下序列：

$$\frac{i - 1/2}{n} \quad i = 1, 2, \cdots, n$$

的所有值（此时 $\delta(n) = \frac{1}{2n}$），见图2—3，a

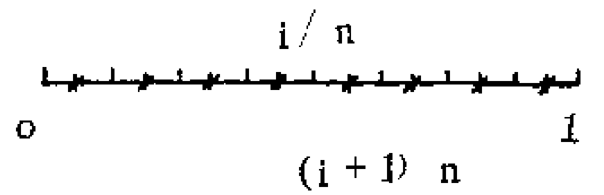

图2—3，a

充分性的证明是平凡的；

对于必要性只需令 $\delta(n) = \frac{1}{2n}$ 将（1）中全部不等式相加便得：

$$1 \leqslant 2n\delta(n) = 1,$$

该式表明（1）中各不等式必须均取等号，这就表明 $\xi'_i = \frac{i - 1/2}{n}$，$i = 1, 2, \cdots, n$。

上述性质是将序列全体作为子样而研究的其均匀性问

题。但是，严格地应研究各部分子样的均匀性问题，例$\delta(k)$，$k\leqslant n$，这个问题是比较困难的。

（3） 伪随机数的独立性

确定伪随机数序列$\xi_1,\xi_2,\cdots,\xi_n$的独立性程度比确定其均匀性程度要困难。对于任意的$x,y\in[0,1]$，令$N_n(x,y)$表示(ξ_1,ξ_2)，(ξ_2,ξ_3)，$\cdots$，(ξ_i,ξ_{i+1})，$\cdots$，(ξ_n,ξ_{n+1})中适合不等式

$$\xi_i<x,\quad \xi_{i+1}<y$$

的个数。根据随机变量间相互独立的定义和频率近似概率的思想，以量

$$\varepsilon(n)=\sup_{0\leq x,y\leq 1}\left|\frac{N_n(x,y)}{n}-\frac{N_n(x)}{n}\cdot\frac{N_n(y)}{n}\right|$$

标志伪随机数序列$\xi_1,\xi_2,\cdots,\xi_n$的独立程度，简称为**独立偏度**。对于固定的n，$\varepsilon(n)$的值越接近零，伪随机数的独立性越好。

还可类似地引进二维问题的均匀性指标：

$$\Delta(n)=\sup_{0\leq x,y\leq 1}\left|\frac{N_n(x,y)}{n}-xy\right|$$

并且显然有下列关系：

$$\Delta(n)\leqslant\sup_{0\leq x,y\leq 1}\left|\frac{N_n(x,y)}{n}-\frac{N_n(x)}{n}\cdot\frac{N_y(y)}{n}\right|$$

$$+\sup_{0\leq x,y\leq 1}\left|\frac{N_n(x)}{n}\cdot\frac{N_n(y)}{n}-xy\right|$$

$$\leqslant\varepsilon(n)+2\delta(n)$$

类似地有

$$\varepsilon(n) \leqslant \Delta(n) + 2\delta(n)$$

从而得到如下结果

$$|\Delta(n) - \varepsilon(n)| \leqslant 2\delta(n)$$

这一结果表明，在伪随机数本身的均匀性保证的情况下，独立偏度 $\varepsilon(n)$与二维均匀性标志 $\Delta(n)$有密切关系。用 $\varepsilon(n)$标志伪随机数的独立性，比用相关系数标志伪随机数的独立性优越得多。

对任意的 S 维情况的独立性问题，根据概率的基本公式：

$$P\{\xi_1 < x_1, \xi_2 < x_2, \cdots, \xi_s < x_s\}$$
$$= P\{\xi_1 < x_1\} \cdot P\{\xi_2 < x_2\} \cdots P\{\xi_s < x_s\}$$

我们用量

$$\varepsilon_s(n) = \sup_{\substack{0 \le x_i \le 1 \\ i=1,2,\cdots,s}} \left| \frac{N_n(x_1, x_2, \cdots, x_s)}{n} - x_1 x_2 \cdots x_s \right|$$

表示 S 维随机数序列的均匀偏度

与 $S=2$ 的情况完全类似，可以证明

$$\Delta_s(n) \leqslant \varepsilon_s(n) + s\delta(n)$$
$$\varepsilon_s(n) \leqslant \Delta_s(n) + s\delta(n)$$

或者表为

$$|\Delta_s(n) - \varepsilon_s(n)| \leqslant s\delta(n), \quad s = 2, 3, \cdots.$$

§4 产生伪随机数的方法

在本节中，我们要介绍几种在计算机上产生伪随机数的

方法；这些方法主要是根据数学中的数论的基本结果。前面已经说过要用递推公式

$$\xi_{n+1}=T(\xi_n)$$

产生伪随机数，因此问题的关键就在于如何选择函数 T。另外，还有一个经济上的问题需要考虑；因为在用蒙特卡罗方法时，伪随机数要用上百万、上亿或更多次，所以如何用经济的方法产生伪随机数是非常重要的问题。归纳起来，我们应考虑的主要问题是：（1）产生伪随机数的质量；（2）方法的经济性；（3）容量要尽可能大。

一、加同余方法产生伪随机数

设 a 为任何满足 $0<a<1$ 的无理数，于是可以构造序列：

$$x_1,x_2,\cdots,x_k,\cdots,\quad x_k=\{ka\}=\{x_{k-1}+a\}$$

符号$\{x\}$表示取 x 的小数部分，显然 $0<x_k<1\quad k=1,2\cdots$。

现在，我们要证明该序列在区间$(0,1)$上是一致分布的，即对任何 a，b 满足 $0\leqslant a<b\leqslant 1$，有

$$\lim_{n\to\infty}\frac{1}{n}N_n(a,b)=b-a,$$

这里 $N_n(a,b)$表示 $x_1,x_2,\cdots,x_n$ 中属于区间 (a,b) 的个数。这种产生伪随机数列的方法称为**加同余方法。**

定理 对任何无理数 $a\in(0,1)$，序列 $x_1,x_2,\cdots,x_n,\cdots$ 其中 $x_n=\{na\}$在$(0,1)$区间上是一致分布的。

证明 首先考虑下列三角级数的和

$$T_n=\frac{1}{n}\sum_{k=1}^{n}e^{2\pi il\{ka\}}=\frac{1}{n}\sum_{k=1}^{n}e^{2\pi ilka}$$

$$i=\sqrt{-1}\ ,\ l=0,\pm 1,\pm 2,\cdots$$

我们有 $T_0=1$，当 $l\neq 0$ 时

$$|T_l|\leqslant \frac{1}{n}\frac{1}{\sin \pi l a}$$

由于 a 是无理数，所以 la 不为整数，$\sin\pi la\neq 0$，于是推得

$$\lim_{n\to\infty} T_l=0\ ,$$

这就意味着对任何三角多项式

$$P(x)=\sum_l a_l e^{2\pi i l x}$$

有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}P(x_k)=a_0=\int_0^1 P(x)dx$$

现在定义函数 $f(x)$ 如下：

$$f(x)=\begin{cases}1 & x\in(a,b)\\ 0 & x\notin(a,b)\end{cases}$$

再构造两个光滑连续函数 f^+ 和 f^-（见图2—4 a），满足

$$f^-(x)<f(x)<f^+(x)$$

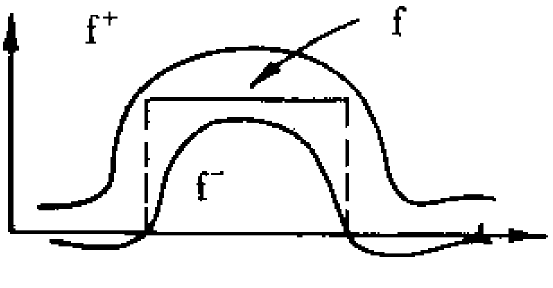

图2—4，a

且有
$$\int_0^1 f^+(x)dx - \int_0^1 f^-(x)dx < \varepsilon$$
其中 $\varepsilon > 0$ 是任意给定的。特别地，不妨假定$f^+(x)$和 $f^-(x)$均为三角多项式，于是有
$$\frac{1}{n}\sum_{k=1}^{n} f^-(x_k) < \frac{1}{n} N_n(a,b) = \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$
$$< \frac{1}{n}\sum_{k=1}^{n} f^+(x_k)$$
所以
$$\int_0^1 f^-(x)dx < \lim_{n\to\infty}\frac{1}{n} N_n(a,b) < \int_0^1 f^+(x)dx$$
进而
$$\int_0^1 f(x)dx - \varepsilon < \lim_{n\to\infty}\frac{1}{n}N_n(a,b) < \int_0^1 f(x)dx + \varepsilon$$
注意到 ε 的任意性以及
$$\int_0^1 f(x)dx = b - a$$
我们推得
$$\lim_{n\to\infty}\frac{1}{n}N_n(a,b) = b - a$$

这样，我们得到一个容量为无穷大的伪随机数序列，可惜由于需要用无理数 a，所以在计算机上是不可能实现的。在计算机中用加同余法时，a 只是某个“很长的”有理小数，数列是近似的。从 $x_{k+1} = \{x_k + a\}$ 可以看出，用该方法产生伪随机数时，每次只作一次加法，因而十分经济。

二、乘同余方法产生伪随机数

按此法产生伪随机数的程序是：对任意初值x_1，由如下递推公式确定

$$x_{k+1} \equiv a x_k \pmod{M},$$
$$\xi_{k+1} = x_{k+1}/M, \quad k = 1, 2, \cdots$$

其中 a 是$(1, M-1)$内的正整数。上式亦可记为

$$\xi_{k+1} = \{a\xi_k\} \quad k = 1, 2, \cdots$$

首先介绍有关整数的若干基本结论：

（1） 同余式$x \equiv y \pmod{M}$是指$x - y$是M的整数倍，即有 $x = y + kM$，k为某整数。

（2） 同余类$(\bmod\ M)$是指用整数M整除后的余数进行分类的各类整数。可以将全体整数分为下列M类：

$$\left.\begin{array}{lll}
0, & \pm M, & \pm 2M, \cdots \\
1, & 1 \pm M, & 1 \pm 2M, \cdots \\
2, & 2 \pm M, & 2 \pm 2M, \cdots \\
\cdots & & \\
r, & r \pm M, & r \pm 2M, \cdots \\
\cdots & & \\
M-1, & (M-1) \pm M, & (M-1) \pm 2M, \cdots
\end{array}\right\} M\text{类}$$

（3） 互素的整数，$(a, b) = 1$表示它们除± 1以外没有任何（整数的）公因子。例如（2，3）= 1，（10，7）= 1；但是（6，3）= 3 ≠ 1，所以整数2，3是互素的，整数10和7也是互素的，但整数6和3不是互素的。

（4） 与模M互素的同余类。在前面的M个同余类中满足$(r,M)=1$的同余类称为与模M互素的同余类。特别是$M=p^\alpha$，p为素数，$\alpha\geq 1$的情况，将所有与模$M=p^\alpha$互素的同余类记为$R=\{r_1,r_2,\cdots,r_\lambda\}$，可以证明$\lambda=p^\alpha-p^{\alpha-1}=p^{\alpha-1}(p-1)$。事实上，$r=p,s$，$s=0,1,\cdots,p^{\alpha-1}-1$时是全部不互素的同余类，所以$\lambda=M-p^{\alpha-1}=p^\alpha-p^{\alpha-1}$。

（5） 若$(a,M)=(a,p^\alpha)=1$和$x\in R$，则$y\equiv ax\in R$. 事实上，由于$(a,p^\alpha)=(x,p^\alpha)=1$和$y=ax+kM$，可立即推出。

（6） 若$(a,M)=1$，$x_1x_2\in R$，$x_1\not\equiv x_2$，则$ax_1\not\equiv ax_2 \pmod M$.事实上，若$ax_1\equiv ax_2$，则$a(x_1-x_2)\equiv 0$，而$(a,M)=1$，所以$x_1\equiv x_2$，于是得到矛盾。

（7） 对任何整数a，满足$(a,M)=1$，则有$a^\lambda\equiv 1 \pmod M$.这里$\lambda=p^\alpha-p^{\alpha-1}$.

事实上，设$y_j\equiv ax_j$，$x_j\in r_j$，$j=1,2,\cdots,\lambda$，于是

$$\prod_{j=1}^{\lambda} y_j\equiv a^\lambda \prod_{j=1}^{\lambda} x_j \pmod M$$

而 $\prod\limits_{j=1}^{\lambda} y_j=\prod\limits_{j=1}^{\lambda} x_j$，所以$(a^\lambda-1)\prod\limits_{j=1}^{\lambda} r_j\equiv 0$，即有$a^\lambda\equiv 1$.

（8） 设μ为使$a^\mu\equiv 1$的最小整数，则μ必整除λ，即$\lambda\mid\mu$.

事实上，设$\lambda\geq\mu$，则$a^\lambda\equiv a^\mu\equiv 1$，即$a^\mu(a^{\lambda-\mu}-1)\equiv 0$和$a^{\lambda-\mu}\equiv 1$；令$\lambda_1=\lambda-\mu$，若$\lambda_1\geq\mu$，同理得$a^{\lambda_1-\mu}\equiv 1$；依次类推必得$\lambda\mid\mu$.

最大容量问题 设$x_1\in R$，$(a,M)=(a,p^\alpha)=1$，对乘同余方法存在最大容量数s，使$x_{s+1}\equiv x_1 \pmod M$，即$(a^s-1)$

$x_1\equiv 0$，于是 $a^s\equiv 1$，可见

$$s\leqslant\lambda=p^a-p^{a-1}.$$

设 $M=2^a$，$x_1\in R$，$a=2k+1$，则 $\lambda=2^{a-1}$．根据 整数的结论 8 ， 5 ， s 应能整除 λ，所以令 $s=2^t$，

$$a^s-1=a^{2^t}-1=(a^{2^{t-1}}+1)(a^{2^{t-2}}+1)\cdots$$
$$(a^2+1)(a^2-1)$$

因为

$$a^2+1=1+1+2^l\cdot 2k+\cdots=2(1+2^l\cdot k+\cdots)$$
$$a^2-1=4(k+1)k=8\times\text{奇数}$$

于是

$$(a^s-1)=2^{t-1+3}\cdot\text{奇数}\equiv 0$$

充要条件为

$$t+2=a,\ \text{即}\quad t=a-2,\ s=2^{a-2}.$$

定理 若 $M=2^a$，$x_1\in R$，a 为奇数，则乘同余法的最容量为2^{a-2}。

乘同余法产生伪随机数列的均匀性

在计算机上我们通常取 $M=2^a$，x_1为奇数。讨论两种情况：

（1） 设 $a\equiv 3\ (\mathrm{mod}\ 8)$，即 $a=8m+3$．伪随机数列 x_i 为：

$$x_1,ax_1,\cdots,a^ix_1,\cdots$$
$$a^ix_i=(8m+3)^ix_1=[(1+2)^i+8*\cdots]x_1$$
$$=(1+2i^2+8*\cdots)x_1,$$

若 i 为偶数，则

$$a^ix_1=x_1+8x_1*$$

对应的伪随机数

$$\xi_i=\frac{1}{2^a}\left(8i+8\left\{\frac{x_1}{8}\right\}\right)\qquad i=0,1,\cdots,2^{a-3}-1;$$

另一串为

$$\eta_i=\frac{1}{2^a}\left(8i+8\left\{\frac{x_2}{8}\right\}\right)\qquad i=0,1,\cdots,2^{a-3}-1.$$

它们是两个等距数列相互交错地排列在（0，1）中（见图2—4，b）

图2—4，b

若 $x_1=2s+1$ 为奇数，则有下列可能性：

$$\left\{\frac{x_1}{8}\right\}=\left\{\frac{1}{8}\right\},\ \left\{\frac{3}{8}\right\},\ \left\{\frac{5}{8}\right\},\ \left\{\frac{7}{8}\right\},$$

$$\left\{\frac{x_2}{8}\right\}=\left\{\frac{3}{8}\right\},\ \left\{\frac{1}{8}\right\},\ \left\{\frac{7}{8}\right\},\ \left\{\frac{5}{8}\right\},$$

$$\delta(n)=\frac{5}{4n},\quad n=2^{a-2}.$$

（2） 若 $a\equiv5(\mathrm{mod}\,8)$，即 $a=8m+5$

$$a^ix_1=(8m+5)^ix_1=(1+4*\cdots)x_1=x_1+4x_1*\cdots$$

$$\xi_j=\frac{1}{2^a}\left(4j+4\left\{\frac{x_1}{4}\right\}\right),\quad j=0,1,2,\cdots 2^{a-2}-1,$$

$$\delta(n)=\frac{3}{4n},\quad n=2^{a-2}.$$

乘同余方法产生伪随机数列的独立性

设 $M=2^a$，则 x_1 为奇数，$a\equiv 5 \pmod 8$ 时，我们得到等距分布的伪随机数 $\xi_1, \xi_2, \cdots \xi_n$。若将其按由小到大排列：

$$\xi_1' \leqslant \xi_2' \leqslant \cdots \leqslant \xi_n'$$

则任意相邻两点距离相等，且

$$\xi_{i+1}' - \xi_i' = \frac{1}{n}, \quad i=1,2,\cdots,n-1.$$

对任意的 $x' < x$，用 $S(x':x)$ 表示区间 (x',x)，$|S(x':x)|$ 表示区间 $S(x':x)$ 的长度。$N(S)$ 表示伪随机数序列 $\xi_1', \xi_2', \cdots, \xi_n'$ 中属于 S 的个数。若伪随机数序列是等分布的，对任意的 $x \in \left[0, 1-\frac{1}{n}\right]$，区间 $S\left(x:x+\frac{1}{n}\right)$ 具有如下性质

$$N(S)=1.$$

完全类似地，对于二维情况，对 $x' < x$，$y' < y$，用 $S(x':x, y':y)$ 表示区域

$$x' \leqslant X < x, \quad y' \leqslant Y < y,$$

$|S(x':x, y':y)|$ 表示区域 $S(x':x, y':y)$ 的面积。$N(S)$ 表示 (ξ_1, ξ_2)，(ξ_2, ξ_3)，$\cdots(\xi_n, \xi_{n+1})$ 中属于 S 的个数。容易看出，只要 $a \leqslant n$，$x \in \left[0, 1-\frac{1}{a}\right]$，$y \in \left[0, 1-\frac{a}{n}\right]$，区域 $S\left(x:x+\frac{1}{a}, y:y+\frac{a}{n}\right)$ 满足 $N(S)=1$.

乘同余方法所产生的伪随机数序列

$$\xi_{k+1}=\{a\xi_k\}=T(\xi_k)$$

的图形如图2—4，c 所示。因 $\left|\xi'_{i+1}-\xi'_i\right|=\frac{1}{n}$， 那么 ξ'_i，ξ'_{i+1} 同属 $T(\xi)$ 的连续范围内，对应的二维点 (ξ'_i,η_i) 和 (ξ'_{i+1},η_{i+1}) 应该有

$$|\eta_{i+1}-\eta_i|=\frac{a}{n}.$$

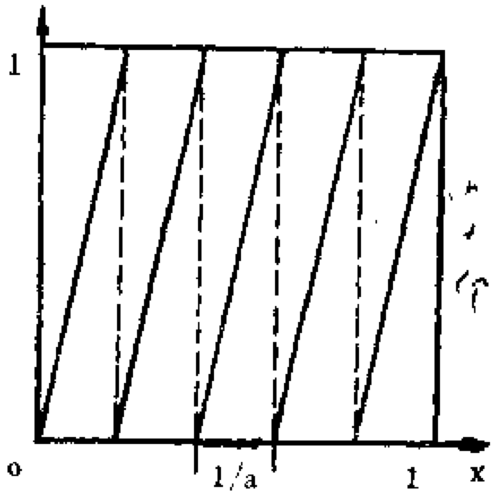

图2—4，c

定理　用乘同余法产生等距伪随机数列，其二维独立性偏度有如下估计：

$$\varepsilon(n)\leqslant\frac{a}{n}+\frac{1}{4a}+\frac{3}{n}.$$

证明　先来计算其均匀偏度 $\Delta(n)$，为此考虑

$$\frac{N_n(x,y)}{n}-xy=\frac{N(S_1)}{n}-|S_1|+\frac{N(S_2)}{n}-|S_2|$$

$$+\frac{N(S_3)}{n}-|S_3|$$

其中（见图2—4，d）

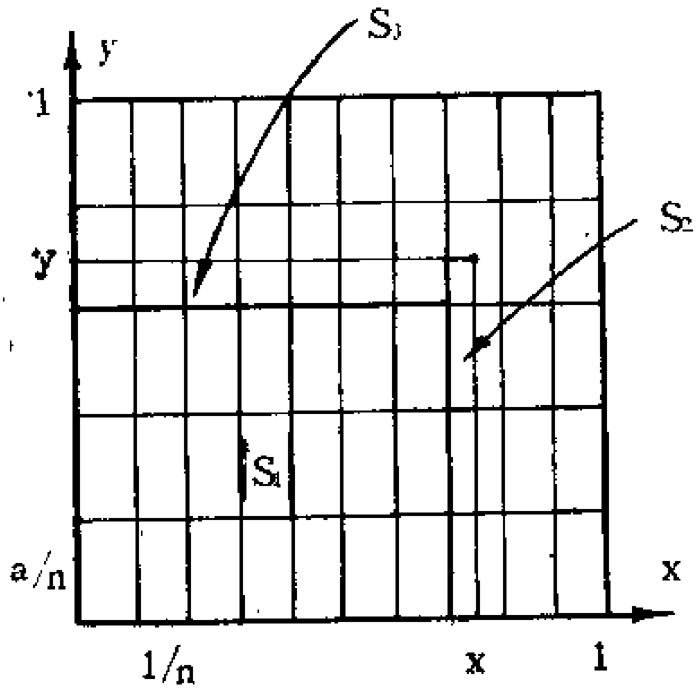

图2—4，d

$$S_1=S\left(0:\left[ax\right]\frac{1}{a},\ 0:\left[\frac{n}{a}y\right]\frac{a}{n}\right),$$

$$S_2=S\left(\left[ax\right]\frac{1}{a}:x,\ 0:y\right),$$

$$S_3=S\left(0:\left[ax\right]\frac{1}{a},\ \left[\frac{n}{a}y\right]\frac{a}{n}:y\right),$$

符号〔·〕表示取其整数部分。

根据前面的讨论，落入每一小矩形格的点（二维点(ξ_i，

ξ_{i+1})）正好只有一个，总共有 $a\times\frac{n}{a}=n$ 个格子，所以首先有

$$\frac{1}{n}N(S_1)-|S_1|=0,$$

其次

$$\frac{N(S_2)}{n}-|S_2|=\begin{cases}\frac{1}{n}\left(\frac{n\{ax\}}{a}+\theta\right)-\frac{\{ax\}}{a}y\\ \qquad 当\ y\geq\{ax\},\\ \frac{1}{n}\left(\frac{ny}{a}+\theta\right)-\frac{\{ax\}}{a}y\\ \qquad 当\ y<\{ax\},\end{cases}$$

其中 $|\theta|\leqslant 1$，于是得到

$$\left|\frac{N(S_2)}{n}-|S_2|\right|\leqslant\frac{1}{n}+\frac{1}{a}\{ax\}(1-y)$$

$$当\ y\geq\{ax\},$$

$$\left|\frac{N(S_2)}{n}-|S_2|\right|\leqslant\frac{1}{n}+\frac{1}{a}(1-\{ax\})y$$

$$当\ y<\{ax\},$$

所以

$$\left|\frac{N(S_2)}{n}-|S_2|\right|\leqslant\frac{1}{n}+\frac{1}{4a};$$

最后考察

$$\frac{N(S_3)}{n}-|S_3|\leqslant\frac{[ax]}{n}-\left\{\frac{ny}{a}\right\}\frac{a}{n}\cdot x\leqslant\frac{a}{n}.$$

总合以上的估计得到

$$\Delta(n)\leqslant\frac{a}{n}+\frac{1}{4a}+\frac{1}{n},$$

$$\varepsilon(n)\leqslant\Delta(n)+2\delta(n)\leqslant\frac{a}{n}+\frac{1}{4a}+\frac{3}{n}$$

$$\left(\delta(n)=\frac{1}{n}\right)$$

而且我们有最佳的 a 值

$$a_{opt}=\sqrt{n}/2,$$

此时

$$\varepsilon_{opt}(n)=\frac{1}{\sqrt{n}}+\frac{3}{n}。$$

三、乘加同余法产生伪随机数

按此法产生伪随机数的程序是，用任意初始值x_1，由如下递推公式确定

$$x_{k+1}\equiv ax_k+C\qquad(\mathrm{mod}\ M)$$
$$\xi_{k+1}=x_{k+1}/M$$

这里 a，c 是属于(1，$M-1$)之间的整数。上式亦可写为

$$\xi_{k+1}=\{a\xi_k+b\}\qquad b=c/M.$$

函数 $T(\xi_k)=\xi_{k+1}$的图形如图 2 — 4， e 所示，相当 于 乘同余法加移位。

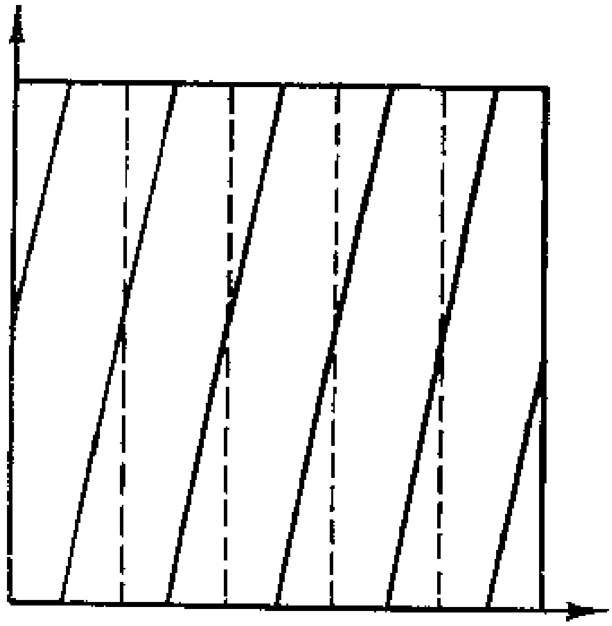

图2—4，e

最大容量问题

设 $M=2^{\alpha}$，$a\equiv 1(mod\ \ 4)$，$c\equiv 1(mod\ \ 2)$，则令

$$x'_{n+1}=a^n x_1+(1+a+\cdots+a^{n-1})c$$

$$=a^n x_1+\frac{(a^n-1)}{a-1}c$$

显然有

$$x_{n+1}\equiv x'_{n+1}=a^n x_1+\frac{(a^n-1)}{a-1}c \qquad (mod\ \ M)$$

若 $x_{n+1}\equiv x_1$，则有

$$(a^n-1)[(a-1)x_1+c]/(a-1)\equiv 0 \qquad (mod\ \ M)$$

注意$[(a-1)x_1+c]$为奇数，所以必定有

$$\frac{a^n-1}{a-1}\equiv 0 \qquad (mod\ M).$$

设 $a=4s+1$，$n=2^a$，则可证明上式成立：

$$\frac{a^{2^a}-1}{a-1}=(a^{2^{a-1}}+1)(a^{2^{a-2}}+1)\cdots$$

$$(a^{2^1}+1)(a^{2^0}+1),$$

$$(a^{2^r}+1)=(4s+1)^{2^r}+1=1+1+4s\cdot 2^r+\cdots$$

$$=2*\text{奇数} \qquad r=0,1,\cdots$$

所以有

$$\frac{a^{2^a}-1}{a-1}=2^a*\text{奇数}.$$

现在设

$$\frac{a^m-1}{a-1}\equiv\frac{a^n-1}{a-1}\equiv 0 \qquad (mod\ M),$$

并且 $m<n$ 是使上式成立的最小整数，则

$$\frac{a^{n-m}-1}{a-1}a^m\equiv 0 \quad \text{即} \quad \frac{a^{n-m}-1}{a-1}\equiv 0,$$

由此，依此类推得到 m 整除 n，因而 m 必为 2^r 的整数，因而 r 必为 a。所以我们证明了乘加同余法的最大容量为 $2^a=M$。

均匀性问题

在 $a\equiv 1(mod\ 4)$，$c\equiv 1(mod\ 2)$，$M=2^a$ 的条件下，x_k 取遍剩余类，因此 $\xi_i'=i/n$，$i=0,1,\cdots,n-1$，$\delta(n)=\frac{1}{n}$.

独立性问题

在上述条件下，可以估计其独立性偏度为

$$\varepsilon(n) \leqslant \frac{a}{n} + \Delta(a,c) + \frac{4}{n}$$

其中

$$\Delta(a,c) = \max\left(\frac{c^2}{4an^2}, \quad \frac{c(n-c)}{an^2}, \quad \frac{(n-c)^2}{4an^2}\right),$$

并且有

$$\frac{4}{25a} \leqslant \Delta(a,c) \leqslant \frac{1}{4a}.$$

当选取 a 为其最佳值

$$a_{opt} = \sqrt{n}/2$$

时，有

$$\varepsilon(n) \leqslant \frac{1}{\sqrt{n}} + \frac{4}{n}.$$

§5 若干应用举例

根据上节所介绍的方法可以在计算机上产生伪随机数。有了伪随机数，我们即可着手解决一些数学问题及物理问题。

一、计算定积分

设我们所要计算的定积分如下：

$$I = \iint_{\Omega} u(x,y)\,dxdy \qquad \Omega = [0,1] \times [0,1].$$

首先，我们需要两个在（0，1）上均匀分布的相互独立的随机数 ξ,η，于是

$$I=Mu(\xi,\eta)$$

即 I 是随机变量 u 的数学期望。现在，我们可取一伪随机数对的序列 (ξ_1,η_1)，(ξ_2,η_2)，…，(ξ_n,η_n)，相应地计算 $u(\xi_1,\eta_1)$，$u(\xi_2,\eta_2)$，…，$u(\xi_n,\eta_n)$。用

$$I_n=\frac{1}{n}\sum_{i=1}^{n}u(\xi_i,\eta_i)$$

作为 I 的近似值。根据柯尔莫哥洛夫定理有

$$P\{I_n-I\xrightarrow[n\to\infty]{}0\}=1,$$

因此 I_n 以概率1收敛于 I。

二、二维随机游动问题（离散的情况）

给定一个网格，如图2—5，a 所示，它由内点 $\{Q,P_k\}$ 和边界点 $\{\Gamma_i\}$ 组成。粒子在任何一个内点以等概率向邻点游动，若到达边界点 Γ_i，则粒子被吸收。对应这个边界点 Γ_i 赋给一常数 $f(\Gamma_i)$。试求由内点 Q 出发的粒子被边界吸收的数学期望 $u(Q)$。

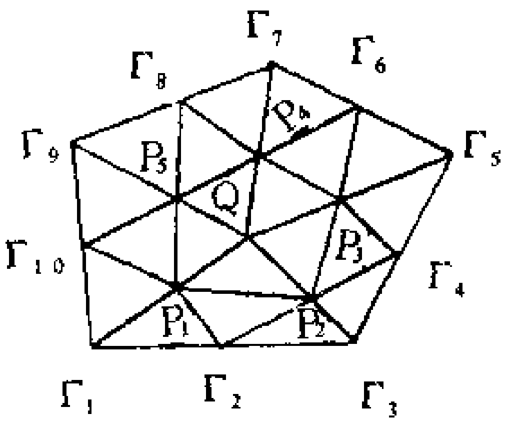

图2—5，a

用$\{\Gamma|Q\}$代表从Q点出发到达边界点Γ的事件，用$\{Q\to P_i\}$代表从Q点出发到达邻点P_i的事件。于是有

$$\{\Gamma|Q\}=\{\Gamma|Q\}\bigcup_{i=1}^{5}\{Q\to P_i\}=\bigcup_{i=1}^{5}\{\Gamma|Q\}\{Q\to P_i\}$$

$$=\bigcup_{i=1}^{5}\{\Gamma|P_i\}\cap\{Q\to P_i\}$$

所以

$$P\{\Gamma|Q\}=\sum_{i=1}^{5}P\{\Gamma|P_i\}P\{Q\to P_i\}$$

因而

$$u(Q)=\sum_k f(\Gamma_k)P\{\Gamma_k|Q\}=\sum_{i=1}^{5}\sum_k f(\Gamma_k)P\{\Gamma_k|P_i\}$$

$$\cdot P\{Q\to P_i\}=\sum_{i=1}^{5}P\{Q\to P_i\}u(P_i)$$

$$=\frac{1}{5}\sum_{i=1}^{5}u(P_i)$$

其中$u(P_i)$表示粒子由内点P_i出发被边界吸收的数学期望。

类似地可推出

$$u(P_1)=\frac{1}{6}\Big[f(P_1)+f(P_2)+f(P_{10})+u(P_2)$$

$$+u(Q)+u(P_5)\Big]$$

如此等等。为了求得$u(Q)$及$u(P_k)$，应该求解这个方程组。

但当网点很多时，这要化费大量的劳动。注意到

$$u(Q)=\sum_{k} f(\Gamma_k)P\{\Gamma_k|Q\}$$

利用它来直接进行抽样：当邻点为5（或6）个时（见图2—5，b），将区间（0，1）等分为5（或6）个小区间

$$I_i=\left[\frac{i-1}{5},\ \frac{i}{5}\right)\qquad i=1,2,\cdots,5$$

（或 $$I_i=\left[\frac{i-1}{6},\ \frac{i}{6}\right)\qquad i=1,2,\cdots,6）$$

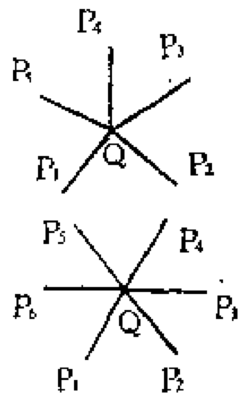

图2—5，b

粒子游动从Q点出发，第一步从一个均匀分布中取一随机数ξ，若$\xi\in I_i$，则表明$Q\to P_i$，即粒子到达了某个点P_i；类似地，再用另一个随机数，确定由P_i到达的下一个邻点。若这个点是某一边界点Γ_j，则游动终止，此时得到$u(Q)$的一个子样$f(\Gamma_j)$，否则继续进行类似的游动…。于是每一次结果可得一个子样，我们记第l次子样为$f(\Gamma_{k_l})$ $l=1,2,\cdots,n$。然后取平均值

$$\bar{u}_n=\frac{1}{n}\sum_{l=1}^{n} f(\Gamma_{k_l})$$

作为$u(Q)$的近似值，用这种方法可以单独计算某个$u(P_i)$，而不必牵涉其他的解$u(P_k)$。根据加强的大数定律，这种近似值以概率1收敛于精确值。

三、解拉普拉斯方程的边值问题（连续的随机游动问题）

在区域 Ω 上解拉普拉斯方程的第一边值问题

$$\begin{cases} \Delta u(x,y)=\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0 \\ u|_\Gamma=g(\Gamma) \end{cases}$$

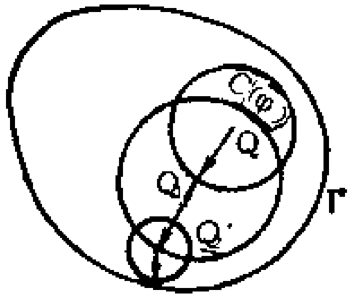

图2—5，c

如图 2—5，c 所示。

现在从内点 Q 出发，以点 Q 为中心作一园 c，$c\subset\Omega$，于是

$$\{\Gamma|Q\}=\{\Gamma|Q\}\bigcup_\phi\{Q\to c(\phi)\}=\bigcup_\phi\{\Gamma|Q\}\{Q\to c(\phi)\}$$

$$=\bigcup_\phi\{\Gamma|c(\phi)\}\{Q\to c(\phi)\}$$

其中 $c(\phi)$表示园周 c 上角参数为 ϕ 的点。因而

$$P\{\Gamma|Q\}=\sum_\phi P\{Q\to c(\phi)\}P\{\Gamma|c(\phi)\}$$

$$u(Q)=\int_\Gamma g(\Gamma)P\{\Gamma|Q\}=\sum_\phi P\{Q\to c(\phi)\}$$

$$\cdot\int_\Gamma g(\Gamma)P\{\Gamma|c(\phi)\}=\int_0^{2\pi}u(c(\phi))\frac{d\phi}{2\pi}$$

$$=\frac{1}{2\pi}\int_0^{2\pi}u(c(\phi))d\phi.$$

所以数学期望 $u(Q)$ 正是该问题的解。这样，我们可以做这个随机游动的抽样：以 Q 点为中心作园 C_Q，使之尽可能地

大，并且不超出Ω（见图2—5，c），然后从Q点向C_Q作随机游动，为此令

$$\phi_1 = 2\pi\xi_1,\ \xi_1 \in (0,1)\text{为一随机数},$$

确定下一个点$Q_1 = c_Q(\phi_1)$，再以Q_1点为中心作园c_{Q_1}，…，直到某个点Q_l，其内接园半径小于某个予先给定的正数$\delta > 0$时，游动终止。取边界Γ上与Q_l距离最近的点Γ_1，得到一个子样$g(\Gamma_1)$，这就完成了一次抽样。然后，再从Q点出发进行新的游动，这样我们得到子样序列

$$g(\Gamma_1),\ g(\Gamma_2),\ \cdots,\ g(\Gamma_N),$$

取其算术平均值

$$\bar{g}_N = \frac{1}{n}\sum_{k=1}^{N} g(\Gamma_k)$$

为$u(Q)$的近似值。

四、发展方程的蒙特卡罗解法

设我们要解一个扩散方程的初值问题：

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},\quad u(x,0) = f(x),\quad x \in (-\infty, +\infty).$$

通常用差分方法将其离散化，令

$$t_k = k\Delta t,\quad x_j = j\Delta x,$$

$$k = 0,1,2,\cdots,\quad j = 0,\pm 1,\pm 2,\cdots,$$

用六点差分方程代替上面的微分方程，得

$$-\frac{r}{2}u_{j+1}^{n+1} + (1+r)u_j^{n+1} - \frac{r}{2}u_{j-1}^{n+1}$$

$$= \frac{r}{2}u_{j+1}^{n} + (1-r)u_j^{n} + \frac{r}{2}u_{j-1}^{n},$$

$$u_j^0 = f(x_j)$$

这里 $r = \frac{\Delta t}{\Delta x^2}$，要求 $r \leqslant 1$，于是上式可以改写成：

$$u_j^{n+1} = \left[\frac{r}{2}(u_{j+1}^{n+1} + u_{j-1}^{n+1}) + \frac{r}{2}(u_{j+1}^n + u_{j-1}^n) \right.$$

$$\left. + (1-r)u_j^n \right] \Big/ (1+r) = \alpha_1 u_{j+1}^{n+1} + \alpha_2 u_{j-1}^{n+1}$$

$$+ \alpha_3 u_{j+1}^n + \alpha_4 u_{j-1}^n + \alpha_5 u_j^n$$

系数 $\alpha_i \geq 0$，$i = 1, 2, \cdots, 5$，$\sum_{i=1}^{5} \alpha_i = 1$。

设想一种随机游动过程，位于点 $(j, n+1)$ 的粒子分别以概率 $\alpha_1, \alpha_2, \cdots, \alpha_5$ 向邻点游动至初始层 $n = 0$ 终止（见图2—5，d），设终止点为 $(x_{k_1}, 0)$，此时赋值以 $f(x_{k_1})$；

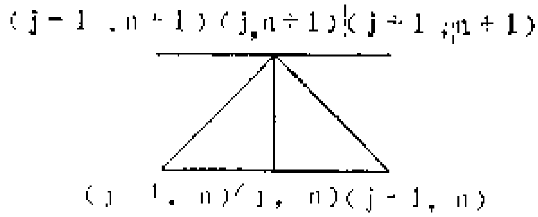

图2—5，d

与离散的二维随机游动问题中的讨论完全类似，如果要想计算 $u(x_0, t_N)$ 的值，为此只需从点 $(0, N)$ 开始作这样的随机游动，直至初始层 $(x_{k_1}, 0)$，于是取得一个 $u(x_0, t_N)$ 的子样 $f(x_{k_1})$；将此过程重复 M 次，我们得到 $f(x_{k_1}), f(x_{k_2}), \cdots,$

$f(x_{k_M})$，于是 $u(x_0, t_N)$ 的无偏估计为

$$\hat{u}_M(x_0, t_N) = \frac{1}{M}\sum_{l=1}^{M} f(x_{k_l}).$$

这种方法可以用来处理很广泛的一类发展方程的初值问题，或初—边值问题。

五、线性代数方程组的蒙特卡罗解法之一

设我们所要求解的线性代数方程组为

$$Ax = f, \quad A = \{a_{ij}\}_{n\times n}, \quad f = \{f_i\}_n,$$

可以将其改写为

$$x = Bx + d, \quad B = \{b_{ij}\}_{n\times n}, \quad d = \{d_i\}_n.$$

假设矩阵 B 的特征值按模小于 1，且 b_{ij} 可写成

$$b_{ij} = b_i p_{ij}, \quad \sum_{j=1}^{n} p_{ij} = 1, \quad p_{ij} \geq 0, \quad j = 1, 2, \cdots, n.$$

因为方程组的解为

$$x = d + Bd + B^2 d + \cdots + B^m d + \cdots$$

所以问题在于如何计算 $B^m d$　$m \geq 1$。为计算 x 的某分量 x_k，对应的 $B^m d$ 的第 k 个分量为：

$$\sum_{i_1}\cdots\sum_{i_m} b_{k, i_1} b_{i_1, i_2} \cdots b_{i_{m-1}, i_m} d_{i_m}$$

或者有

$$\sum_{i_1}\cdots\sum_{i_m} p_{k, i_1} p_{i_1, i_2} \cdots p_{i_{m-1}, i_m}$$
$$\cdot \, (b_k b_{i_1} \cdots b_{i_{m-1}} d_{i_m})$$

其中令