

文章编号: 1007-4627(2005)04-0395-03

## 蒙特卡罗方法的两个基本问题\*

王仲奇<sup>1,2</sup>, 宋玉琳<sup>3</sup>, 肖刚<sup>1</sup>, 张本爱<sup>1</sup>

(1 北京应用物理与计算数学研究所, 北京 100088;

2 中国原子能科学研究院核技术应用研究所, 北京 102413;

3 固原高等师范专科学校, 宁夏 固原 756000)

**摘要:** 随机数的产生和概率分布的随机抽样方法, 是蒙特卡罗方法的两个最基本的问题。对这两个问题的处理目前仍然因循 60 年前蒙特卡罗方法诞生初期的技术途径, 因此长期受一些无法彻底解决的问题所困扰, 同时也面临一些新问题的挑战。分析了上述问题, 分别对随机数的产生和任意分布的抽样方法给出了新的技术解决途径。

**关键词:** 随机数; 伪随机数; 随机抽样方法; 自动抽样方法

**中图分类号:** O21      **文献标识码:** A

### 1 引言

现代蒙特卡罗方法自 20 世纪 40 年代的曼哈顿工程起, 已经有 60 年的研究和应用历史了。在 60 年的发展过程中, 蒙特卡罗方法的应用一方面被拓展到核物理以外的许多科学领域, 另一方面在核工程、核技术、核安全等核物理相关领域中也得到了广泛深入的应用。

蒙特卡罗方法是由数学、概率统计和计算技术交叉结合形成的计算方法, 在此基础上逐渐形成了一套成熟有效的研究和应用体系。任何一个体系一旦形成并相对成熟, 它的体系基础就会由于完整而失去进一步发展的内在动力, 而与体系的外部发展、与对体系的需求发展相脱节。基于这样的考虑, 我们需要适时地重新审视和不断地发展蒙特卡罗方法的体系基础, 及时吸收最新思想和最新成果, 以适应对蒙特卡罗方法的需求。

随机数产生方法和随机抽样方法是蒙特卡罗体系的两个最基础、最重要和最成熟, 但又同时是几十年来发展最不明显的组成部分。本文将分析研究目前常规的随机数产生方法和随机抽样方法的由来和存在的问题, 给出常规方法的替代方法。

### 2 随机数的产生方法

在蒙特卡罗方法中, 为实现对随机对象的模拟, 需要利用随机数( $(0, 1)$ 中均匀分布的随机变量)来构造所需分布的随机变量。随机数产生方法的优劣将直接影响蒙特卡罗模拟研究的成败。从原理上看, 我们在蒙特卡罗方法中所需要的随机数应该是一个完全没有记忆、完全不可压缩的、在 $(0, 1)$ 中各处几率相同的数的无穷数列。然而, 这样的数列在产生、保存和使用上存在许多问题, 我们必须降低对随机数的要求, 尤其是在几十年前蒙特卡罗方法产生的初期。

#### 2.1 “真”随机数(表)<sup>[1]</sup>

几乎作为一种本能, 人们会使用掷骰子和掷硬币的方式来产生随机数。随着技术的发展, 噪声和放射性衰变也被利用作为产生随机数的方法。然而, 由于蒙特卡罗方法所需要的随机数的数目是非常巨大的, 这就带来一系列的问题, 如产生成本、保存随机数(表)对存在介质的需求等。这些问题在 60 年前是非常严重的本质性问题, 是蒙特卡罗方法可行性的问题。为了解决这一系列问题, 人们提出了伪随机数的概念, 回避了上述问题, 使蒙特卡罗

收稿日期: 2005 - 09 - 15

\* 基金项目: 计算物理实验室资助项目

作者简介: 王仲奇(1962-), 男(汉族), 四川巴塘人, 研究员, 博士, 从事蒙特卡罗方法理论与应用研究;

E-mail: zqwang62@gmail.com

方法得以发挥巨大作用。

### 2.2 伪随机数及其存在的问题

伪随机数(列)是利用确定的递推公式得到的、在一定程度上体现随机性的数列。由于伪随机数与我们希望的随机数之间存在本质的区别,引入伪随机数会给蒙特卡罗方法带来一系列的问题,如均匀性、独立性和周期性等经典问题<sup>[1]</sup>。随着对并行计算要求的扩展,伪随机数的并行问题日益突出。上述问题从总体上讲是数学问题。而 Landau 给出的是从物理上对伪随机数的批判,“好”的随机数产生差的物理结果<sup>[2]</sup>。这迫使人们又回到 60 年几乎同样的水平来看待蒙特卡罗方法的可行性问题,所不同的是与 60 年前相比,可以依赖的技术发生了巨大的变化,人们的处境变得不那么尴尬了。

### 2.3 对“真”随机数的回归

随着技术的发展,有了更多有效且廉价的产生随机数的方法,有了非常廉价的计算机存储介质。这使得人们有可能在一个高的技术起点上更好地解决 60 年前的问题,从而在原理上彻底解决困扰人们 60 年之久的伪随机数的均匀性、独立性和周期性的问题。解决了日益突出的可并行问题,最重要的是解决了计算的可靠性问题。而需要做的是购买更大的存储空间和下载随机数<sup>[3]</sup>。

## 3 随机抽样方法

在蒙特卡罗方法中,随机抽样是在计算机上构造研究对象的随机性质的关键技术,而随机抽样方法是实现蒙特卡罗方法的核心问题。与随机数的产生问题一样,人们需要回到 60 年前。不是任意一个概率分布都有确定的抽样方法,不是所有确定的抽样方法都便于使用。挑选法<sup>[4]</sup>非常好地解决了上述问题。人们可以利用挑选法实现对任意分布的抽样,所付出的代价是降低抽样效率(本质上是增加了抽样所需计算机时间和增加对随机数的耗费)、增加了抽样的时间不确定性(这严重影响随机抽样的并行化,而根本上影响了蒙特卡罗的并行化)和抽样方法严重地依赖于具体分布(严重影响了蒙特卡罗程序的智能化和傻瓜化)。人们迫切希望解决这一现状。

### 3.1 基于 alias 思路的抽样自动化

A. J. Walker 在 1974 年给出了一个被称为

alias 的算法<sup>[5-7]</sup>,非常巧妙地优化了对离散性随机变量的抽样,其最主要的特点是消除了抽样的时间不确定性和抽样的分布依赖性,在此基础上人们可以构造对任意分布的自动抽样算法。这在一定程度上实现了随机分布的抽样自动化。然而基于 alias 的抽样自动化在分布依赖上,尤其是对于无限分布的抽样,还存在一些问题。

### 3.2 基于偏倚思路的抽样自动化

对于蒙特卡罗模拟来说,我们可以而且在很多场合希望使用带权重的样本。因此构造的自动抽样方法可以不受囿于 alias 算法的限制,而充分利用蒙特卡罗方法广泛应用的偏倚抽样思想来构造任意分布的自动抽样方法。

由蒙特卡罗偏倚抽样方法我们知道,在积分问题中,在适当的条件下,分布  $f(x)$  的样本可以用对(偏倚)分布  $g(x)$  的带有权重  $f(x_f)/g(x_f)$  的样本来替代。基于这样的思想,根据不同的要求,我们可以选取相对固定的偏倚分布(比如均匀分布对于有限分布,指数分布对于无限分布),通过这些分布的抽样和相应的纠偏来实现对任意分布的抽样。

图 1 给出了基于偏倚抽样思路的对指数分布  $f(x)=\exp(-x)$  的抽样结果,所用的偏倚分布为  $g(x)=0.005\exp(-0.005x)$ 。从图中可以看出:首先与常规指数分布抽样方法相比其样本可从区域得到很大的拓展,这对于解决屏蔽计算的深穿透问题很有帮助;其次在一个很大的范围里,样本分布与总体分布的偏差稳定在一个水平上,同时样本数的分布非常近似于常数分布,这些特点对于解决蒙特卡罗整体计算的误差均衡问题很有帮助。

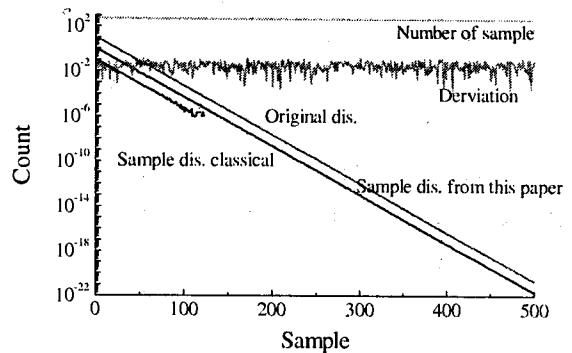


图 1 指数分布的指数偏倚抽样

图中的样本指数分布、总体指数分布和常规抽样的指数分布是重合的,为了方便区分,各自乘了一个常数。

本文给出了关于随机数产生和任意分布抽样的新的解决方法和技术途径,它们不仅可以作为目前常规方法的补充,而且对于一些蒙特卡罗方法长期

未解决的重要问题,它们也是一条很好的解决途径。我们从目前的研究中已经看到它们很好的应用前景<sup>[8]</sup>。

### 参 考 文 献:

- [1] 裴鹿成,张孝泽. 蒙特卡罗方法及其在粒子输运问题中的应用. 北京: 科学出版社, 1980.
- [2] Ferrenberg A M, Landau D R, Wang Y J. Phys Rev Lett. 1992, **69**(23): 3 382.
- [3] <http://www.random.org/> 和 <http://www.randomnumbers.info/>.
- [4] Fishman G S. Monte Carlo: Concepts, Algorithms, and Applications, Springer, 1995.
- [5] Walker A J. Electronic Letters, 1974, **10**: 127.
- [6] Walker A J. ACM Trans Math Software, 1977, **3**: 253.
- [7] 王仲奇,肖刚,张本爱. 随机抽样的优化研究. 计算物理, 待发表(已接收).
- [8] 王仲奇,肖刚,张本爱. 通量场精细计算的蒙特卡罗技巧. 计算物理, 待发表(已接收).

## Two Foundational Problems in Monte Carlo Method\*

WANG Zhong-qi<sup>1,2</sup>, SONG Yu-lin<sup>3</sup>, XIAO Gang<sup>1</sup>, ZHANG Ben-ai<sup>1</sup>

(1 Lab. Com. Phys., Institute of Applied Physics and Computational Mathematics, Beijing 100088, China;

2 Dept. of Nuclear Tech., China Institute of Atomic Energy, Beijing 102413, China;

3 Guyuan Teachers' Colleague, Guyuan 756000, Lingxia, China)

**Abstract:** The production of random number and the sampling of random distribution are two foundational problems in Monte Carlo Method. In this paper, we carry out some new ways to treat the problems.

**Key words:** random number; pseudo random number; sampling; auto-sampling

\* Foundation item: Laboratory Foundation of Compute Physics