蒙特卡洛模型在矿产资源量预测中的应用

朱海宾
(新疆大学 地质与勘查工程学院, 乌鲁木齐 830046)

摘 要: 蒙特卡洛法是根据统计抽样理论, 通过对随机变量函数的概率模拟、统计试验来进行近似求解的方法。矿产资源无论是作为地质过程的产物还是作为地质观测结果都具有随机性, 对资源量的估算一般受概率法则的支配, 因而是一定概率意义下的估算。蒙特卡洛方法能够正确地模拟随机变量的分布, 再现其取值规律。文章结合实例阐述了蒙特卡洛模型在矿产资源量预测中的应用。

关键词: 蒙特卡洛法; 储量估计; 预测模拟; 概率统计
中图分类号: P628; P612 文献标识码: A 文章编号: 1001-1412(2010)01-0050-05

0 引言

蒙特卡洛法(Monte-Carlo method)又称概率统计法, 它是以随机变量为对象, 以概率论为理论基础, 通过对随机变量的概率模拟、统计试验来近似求解的一种方法。应用蒙特卡洛方法对资源量进行预测, 是按照随机抽样理论对资源量的分布进行模拟的过程, 计算得到的结果是一条储量概率分布曲线。从统计观点看, 矿体形成的实质是对所研究地质体的抽样观测过程。我们把直接观测到的地质数据看作基础地质信息, 而把经统计方法得到的统计量看作是复合地质信息。以后随着工作的增加、数据的积累, 才能得到对地质体的无偏估计(即一定概率下的预测资源量)。蒙特卡洛方法就是依据统计预测方法建立的数学模型。

1 应用蒙特卡洛方法模拟资源量

应用蒙特卡洛模型进行资源量预测的基本思想是: 为了求解地质问题, 首先根据已有数据建立一个符合地质过程的概率模型, 然后通过对模型的抽样试验来计算所求参数的统计特征, 最后求出具有特定期望值的近似解。

蒙特卡洛模拟计算过程为: 首先建立概率模型, 然后产生符合要求的抽样随机数, 最后用随机数对已知分布的地质随机参数进行抽样, 计算待求解。

资源量包括质量和数量两方面特征, 属于前者的有品位, 属于后者的有矿床数、矿石量及金属量等, 以下把它们统称之为参数, 资源量就用这些参数来表达。

2 建立概率模型

蒙特卡洛方法的应用非常广泛, 只要能构建出适当的概率模型, 几乎可以用来模拟任何问题。因此, 构建概率模型是关键的一步。当我们研究的地质对象是按概率法则变化时, 在数学模型中加入满足概率法则的随机数来反映这一随机变化的过程, 由此所建立的模型就是随机模型。目前在矿产资源总量预测中, 常用的几种模型是: ①随机变量乘积模型; ②随机变量和模型; ③随机变量混合模型。这里的随机变量都是指与资源量有关的参数, 建立模型就是在各个参数分布已知的条件下, 通过参数之间的不同关系, 求得潜在资源量的分布规律。不同的模型反映不同性质的问题, 因此有不同的适用范围。

下面的实例比较简单, 仅涉及乘积概率模型, 通过它探讨构建概率模型的方式和参数的使用条件。

收稿日期: 2009-02-27; 改回日期: 2009-03-31
基金项目: 国家重点基础研究发展计划项目(2007CB411308)资助。
作者简介: 朱海宾(1984-), 男, 山东荷泽人, 硕士研究生, 主要研究方向为矿产资源预测评价。通信地址: 新疆乌鲁木齐市延安路 1230 号新疆大学南区地质与勘查工程学院; 邮政编码: 830049; E-mail: zhudahai1984@163.com

实例: 某种与基性- 超基性岩有关的矿产资源, 在某个区域内有 29 个已知矿床, 将它们的矿石量、品位数据列入表 1。此外, 还知道矿区外围基性- 超基性岩体的分布情况, 现在把这些未知岩体作为研究对象, 来估算预测区(即已知矿区外) 的资源潜力。

表 **1**　矿石量及品位数据表

Table 1　Ore tonnage and grade

矿床序号	矿石量(10^4t)	品位(%)
1	270213. 34	2. 19
2	250060. 67	0. 59
3	130523. 42	0. 63
4	35100. 26	0. 83
5	160002. 63	0. 64
6	49433. 93	0. 26
7	50204. 46	0. 59
8	12101. 3	0. 73
9	14921. 72	0. 27
10	32012. 39	0. 42
11	17031. 45	0. 38
12	13581. 68	0. 37
13	10451. 79	0. 3
14	8856. 72	0. 7
15	8642. 52	0. 4
16	1630. 42	1. 03
17	1302. 58	0. 38
18	4429. 92	0. 7
19	1501. 42	0. 5
20	210. 1	0. 33
21	490. 55	0. 32
22	4432. 87	0. 31
23	58. 53	0. 25
24	1865. 43	0. 25
25	52000. 49	0. 25
26	49500. 32	0. 39
27	4299. 43	0. 24
28	16003. 61	0. 5
29	8000. 39	0. 25

显然, 这个问题可通过已知矿床建立资源量概率模型, 然后把它推广到预测区来解决。用金属量来表示资源, 建立概率模型 $M= T \times C$, 这里 M 是一个随机变量。对于模型区的 29 个矿床来说, M 不再是已知储量。由于 M 是根据已知矿床的参数 T 和 C 随机相乘得到, 它的规律性表现在随机变量 M 的分布上。当把这个模型推广到预测区时, 相当于假设了一个前提: 预测区内"矿床"的资源量也服从 M 的分布。由于待测岩体并非全部都是矿床, 不能直接带入模型, 因此需要通过某种方法在预测区对各个岩体给出一个有利性指标, 把预测对象调整为"矿床"的概念, 使之符合模型的使用条件。根据这种想法, 可将前面的模型改为: $M= T \times C \times L$, 其中 L 是(0, 1) 区间上的预测样品的成矿有利系数。

构建概率模型的方法很多, 根据问题的性质和资料的水平不同, 可以有不同的模型。上面是通过对各个矿床的研究来求整个预测区的资源量, 实际上还可以从其他途径入手, 通过各个已知矿田之间关系的研究来求预测区资源量。

3　资源参数分布的模拟

在资源预测阶段, 所能收集到的地质参数数量一般比较少, 构建随机变量的分布函数时, 除人们直接给出的资源参数分布以外, 一般由实测数据建立统计分布, 有两种方法: 一是选用合适的已知分布律来拟合, 二是用数学方法构造分布函数。

3. 1　选用合适的已知分布律拟构建分布函数

对参数的原始数据进行格理, 分组求频数, 作出频率直方图, 根据直方图的峰度、偏度等特征, 选用已知的分布律来代表参数的分布。在实际工作中, 我们可以根据直方图的形态用多种分布来试验, 选择拟合程度较好的来使用。用已知分布律来代表资源参数的分布, 给我们研究问题带来很大方便。

以表 1 的 29 个已知矿床的储量资料为例。将矿石量数据取对数, 用 0. 501 为组距, 分为 9 组, 作出频数表(表 2) 并绘制相应的频率直方图(图 1) 。

根据直方图单峰、对称的特征, 用正态分布来拟合它。为此求得(对数) 算术平均值 α= 4. 001, 标准差 σ = 0. 875。在表 3 中给出实测频率和参数(α= 4. 001, σ= 0. 875) 的正态分布理论概率。从表 3 中可以看出理论概率和实测频率很接近, 故认为用正

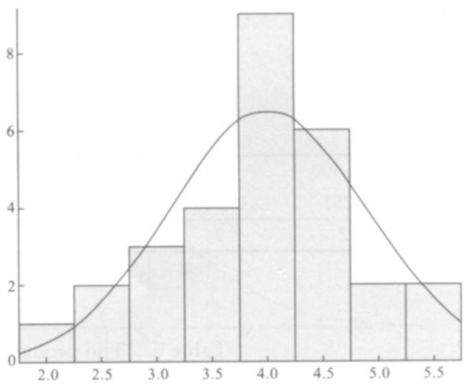

图 **1**　矿石量频率直方图

Fig. 1　Ore tonnage histogram

态分布来拟合直方图是合适的。严格地讲,理论分布曲线的选配需要经检验,只有在理论分布与实测结果没有显著差异时才使用。

表 **2** 矿石量频数表

Table 2 Frequency of ore tonnage

组中值									
	对数	2.003	2.496	3.002	3.496	4.003	4.498	5.004	5.503
	真数	100.69	313.33	1004.62	3133.29	10069.32	31477.48	100925.29	318419.75
频数		1	2	3	4	9	6	2	2

表 **3** 正态分布拟合矿石量频率表

Table 3 Frequency of normally simulated ore tonnage

组中值(对数)	2.003	2.496	3.002	3.496	4.003	4.498	5.004	5.503
实测频率	0.034	0.034	0.069	0.103	0.172	0.241	0.069	0.069
正态理论频率	0.007	0.021	0.054	0.111	0.171	0.208	0.129	0.070

3.2 用数学方法构建分布函数

构建分布函数 $F(x)$ 也是在建立参数的频率直方图基础上进行。具体作法是寻找合适的函数 $f(x)$,用它来拟合频率直方图,$f(x)$ 须满足作为密度函数的条件,即:$f(x) \geqslant 0$, $\int_{-\infty}^{+\infty} f(x)\,dx = 1$。这是一个曲线拟合的问题,可以用样条函数逼近直方图来构建 $f(x)$,有了 $f(x)$ 自然就可以得到 $F(x)$。二次样条函数的计算公式为 $f(x) = \sum_{j=0}^{N} y_i \Omega\left(\frac{x - x_j}{h}\right)$,其中 y_i 为直方图每个柱的高度(频数),h 为柱的宽度(组距),$\Omega(t)$ 为基本样条函数。

下面以 29 个矿床的品位为例,说明 $F(x)$ 的求法。将表 1 中品位数据取对数,以 0.25 为组距分为 9 组,作出统计频数表(表 4)及直方图(图 2)。由直方图看出,原始数据取对数后图形仍不对称,故不能用对数正态分布来描述品位的分布状态。我们采用二次样条函数来拟合,利用公式可求出任意点的概率密度,积分即得到任意区间的概率。为便于比较,将算得的各组理论概率列在表 5 中。

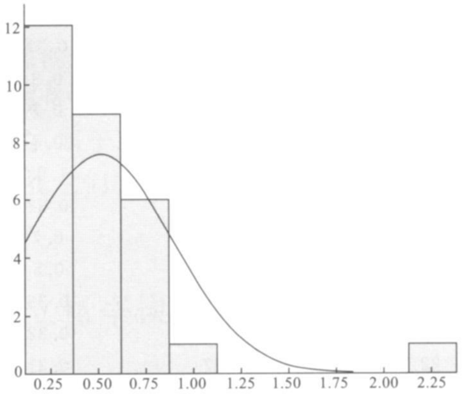

图 **2** 品位频率直方图

Fig.2 Ore grade frequency histogram

构建函数 $f(x)$ 来拟合直方图的方法优点是算法统一,便于编制程序,而且应用中无须事先知道随机变量服从何种分布。

表 **4** 品位频数表

Table 4 Frequency of ore grade

组中值										
	对数	−0.600	−0.348	−0.124	0.002	0.096	0.175	0.243	0.301	0.352
	真数	0.251	0.449	0.752	1.004	1.248	1.496	1.750	2.00	2.25
频数		12	9	6	1	0	0	0	0	1

表 **5** 样条函数拟合品位概率表

Table 5 The probability of ore grade spline function simulation

分组	1	2	3	4	5	6	7	8	9	总计
实际频率	0.413	0.310	0.207	0.034	0	0	0	0	0.034	1
理论频率	0.415	0.302	0.213	0.031	0.001	0.003	0.006	0.007	0.028	1

4　随机数与抽样

随机数就是各种不同分布随机变量的抽样序列, 在序列内部无任何规律可寻, 随机数的产生, 要求分布的均匀性、抽样的随机性、试验的独立性、前后的一致性等。

4.1　随机数的产生

产生随机数的方法有多种, 在蒙特卡洛方法抽样模拟中所使用的随机数是由计算机上用某种数学方法计算出来的。由于它受计算机字长的限制而具有一定周期, 因此不是真正的随机数, 这样产生的随机数被称作“伪随机数”。在应用中, 可选择较好的算法使“伪随机数”有尽可能长的周期来保证抽样的随机性。在计算机上产生“伪随机数”经常用的方法是乘同余法, 计算公式为 $x_{n+1}=\lambda x_n(\mathrm{mod}M)$, 这是一个递推公式, 给定 x_0 以后就可以逐个计算 x_1, x_2, …; 譬如已计算了 x_n, 那么将 x_n 乘以常数 λ, 用 M 来除, 取其余数即为 x_{n+1}, 例如取 $M=59$, $\lambda=2$, $x_0=14$, 算得随机数为: 14 28 56 53 47 35 11 22 44 29 58 57 55 51 43 27 54 49 39 19 38 17 34 9 18 36 13 26 52 45 31 3 6 12 24 48 37 15 30 1 2 4 8 16 32 5 10 20 40 21 42 25 50 41 23 26 33 7。这是一个整随机数列, 它在(1, 58)上均匀分布, 要想把它变换到(0, 1)上去, 只要将每个数减去 1 再除以 57 即可。在抽样过程中使用的就是(0, 1)上均匀分布的随机数, 当然这里面不限于 58 个数了。

4.2　抽样模拟

有了随机数后, 就可进行抽样了。所谓抽样, 就是在已知的某个随机变量分布情况下, 通过取随机数实现在该变量中一次次取值的过程。例如准备对品位模型 C 的分布进行抽样, 为此将 C 的取值区间分成 n 份, 则品位落在各个小区间上为事件 C_1, C_2, ……C_n, 而相应的概率为 p_1, p_2, ……p_n。若令 $p^k=\sum_{i=1}^{K}p_i$ 和 $p_0=0$, 则 p^k 表示了前 k 个概率之和, 显然有 $p^k-p^{k-1}=p_k$。现取(0, 1)上均分分布的随机数 r, 若 r 落在(p^{k-1}, p^k)上, 则说事件 C_k 发生(即取得了一个品位为 $C=C_k$ 的样品), 其概率为 p_k, 这样, 就使一个随机数 r 与 C 的一个取值 C_k 对应起来, 从而完成一次抽样。

抽样方式不同, 随机结果也不相同。抽样方法的使用取决于 2 个条件: 一是概率模型的形式, 二是参数的性质。仍以岩体的预测模型 $M=T\times C\times L$ 为例, 式中参数 T 和 C 的分布已经在前面给出, 又假设类比系数 L 的分布也已做出, 那么对每个预测岩体金属储量的估计值可由下述抽样过程模拟出来: 先取一个(0, 1)上均匀分布的随机数 r_1, 在 T 的分布中抽得一个矿石量 t_1, 然后再取随机数 r_2, 在 C 中抽得一个品位值 c_1, 第三次取随机数 r_3, 在 L 中抽得 l_1, 将这 3 个值相乘, 便得到一个预测岩体金属量的随机值 $m_1=t_1\times c_1\times l_1$, 于是完成了一轮抽样。依此做下去, 譬如进行了 1 000 轮抽样, 便有 1 000 个金属量的取值 $m_1, m_2, \cdots, m_{1000}$, 对这 1 000 个数分组求频, 绘制直方图, 根据直方图的峰度、偏度等特征值, 选用已知统计分布拟合可得潜在资源量的分布, 也可根据抽样结果利用数学方法直接构建潜在资源量分布函数。

5　资源量的估算

一系列随机抽样的结果, 得到一系列资源量的取值, 对这些值进行统计整理便可得到资源量的概率分布 $p(x)$ 和概率分布函数 $F(x)$, x 代表资源量, 或者是矿石量、品位等。一般资源量的表达都使用 $F(x)$, 它实际上就是累积概率。在估算矿产资源储量时, 总是希望知道在 100% 概率下的资源量有多少, 所以在储量资源评价中把随机变量的分布函数定义为 $\overline{F}(x)=1-F(x)=1-p(\xi<x)=p(\xi\geqslant x)$, 这是定义在(0, +∞)上的单调减函数, 其图形如图 3。

根据这条曲线就可以估出计任何概率下的资源量。例如横坐标代表资源量, 纵轴代表累积概率, 则 y 轴 0.75 对应的 x_1 代表 75% 的概率下储量不超过 x_1 吨。这样就可以用该数值进行研究区内潜在资

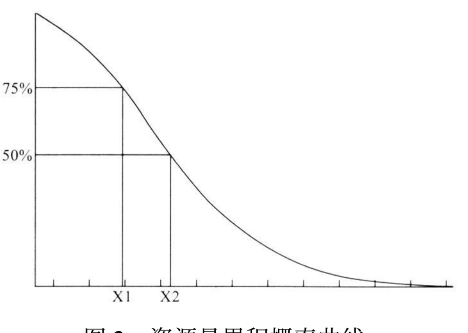

图 3　资源量累积概率曲线

Fig. 3　Curve of cumulative probability of mineral resource volume

源量的评价。

6 蒙特卡洛模型的应用

使用蒙特卡罗方法对资源量的预测过程是应用随机抽样的理论对资源量的分布进行模拟的方法来完成的。它所依据的原始资料主要是矿床数、矿石量、品位及储量等数据。这些数据可以来源于实际观测,也可以由人根据经验估计给出。作为一种计算技术,对于前者,它属于矿床模型法;而对于后者,则属于主观概率法。不论哪种方法,都须注意一个资源含义的问题。由于以往的工作基本上是研究“储量”,而对资源预测来说,这些储量数据是不够的。因为依靠储量资料所做出的估计并非全部资源,它没有包括“潜在”的那一部分在内。通常这类资料不易直接得到,它们很多未经整理,分散在各种原始资料中。由于过去对这一部分资源的特点和规律注意得不够,因此即使是专家的估计也可能出现偏差,所以应当有目的地搜集和整理这些资料,在计算中加以利用。

金属资源量的估算是个概率数字,在找矿工作中有一定的参考价值,但是矿产资源的普查工作应该建立在地球化学、地球物理等数据分析的基础上,重点对异常区域进行普查,估计区域矿产资源潜力对勘查未探明矿产资源起着指导作用。

参考文献:

[1] 朱裕生. 矿产资源评价方法导论[M]. 北京: 地质出版社, 1984.

[2] 杨本锦, 郭履和. 矿产资源评价决策模型[M]. 成都: 四川科学技术出版社, 1985.

[3] 赵鹏大, 胡旺亮. 矿床统计预测[M]. 北京: 地质出版社, 1994.

[4] 侯景儒, 郭光裕. 矿床统计预测及地质统计学的理论与应用[M]. 北京: 冶金工业出版社, 1993.

[5] 侯景儒, 黄竞先. 地质统计学及其在矿产储量计算中的应用[M]. 北京: 地质出版社, 1982.

[6] 裴鹿成, 王仲奇. 蒙特卡洛方法及其应用[M]. 北京: 海洋出版社, 1998.

[7] 王凤清, 刘颖, 崔军. 蒙特卡洛模拟法在辽宁层控碳酸盐岩型铅锌银矿产预测中的应用[J]. 辽宁地质, 1992, (6): 175-186.

[8] 李琼. 用蒙特卡洛方法进行油气资源经济评价[J]. 国外油气地质信息, 2003, (3): 36-39.

[9] 宋国耀, 张晓华, 朱裕生. 矿产资源潜力评价的若干问题[J]. 中国地质, 1999, (8): 17-19.

[10] 文环明, 肖慈珣. 蒙特卡洛法在油气储量估算中的应用[J]. 成都理工学院学报, 2002, 10(3): 21-23.

[11] 王忠法. 应用蒙特卡洛方法进行预测的探讨[J]. 方法研究, 1990, 9(5): 51-54.

[12] 罗汉, 彭国强. 概率论与数理统计[M]. 北京: 科学出版社, 2007.

[13] 柳海东. 蒙特卡洛方法在概率计算中的应用[J]. 苏州职业大学学报, 2004, 12(8): 69-70.

[14] 张寿庭, 赵鹏大. 多目标矿产预测评价及其研究意义[J]. 成都理工大学学报, 2003, 30(5): 441-446.

[15] 杨永华. 蒙特卡洛法与矿产资源评价[J]. 地质与勘探, 1985, 23(1): 45-50.

[16] 赵鹏大, 张建平, 张廷寿. “三联式"成矿预测新进展[J]. 地学前沿, 2003, 10(2): 455-463.

[17] 赵玉琛. 成矿预测中非参数蒙特卡洛模拟方法研究[J]. 矿产与地质, 1993, (5): 191-199.

[18] 王银宏, 严光生, 翟裕生. 三部式潜在矿产资源定量评价与蒙特卡罗模拟[J]. 中国矿业, 2006, 11(6): 50-53.

THE APPLICATION OF MONTE-CARLO MODEL IN THE ASSESSMENT OF MINERAL RESERVES

ZHU Hai-bin

(Geology and Exploration Engineering Department, Xinjiang University, Urumqi 830046, China)

Abstract: Monte-Carlo method is derived from statistically sampling theory with solution acquired approximately by probability simulation of the random varible function and statistical tests. As a matter of fact, whether mineral resource is a product of the geological process or that of geological observation is all random. Generally, the mineral resource assessment is governed by probability principle and is of probability sense. Monte-Carlo method can simulate accurately the distribution of random varibles and reappears sampling regularity. This paper expatiates the application of Monte-Carlo method in the assessment of mineral reserves combined with examples.

Key Words: Monte-Carlo method; reserves estimation; forcast simulation; probabilistic statistical method