数学地质进展 2

矿产储量的地质统计学评价

M·戴维 著

地质出版社

矿产储量的地质统计学评价

数 学 地 质 进 展 2

矿产储量的地质统计学评价

(加拿大)M·戴维 著

孙惠文 刘承祚 译

地 质 出 版 社

Developments in Geomathematics 2

GEOSTATISTICAL ORE RESERVE ESTIMATION

by Michel David

1977

数学地质进展 2

矿产储量的地质统计学评价

(加拿大)M·戴维 著
孙惠文 刘承祚 译

责任编辑：杨友爱
地质出版社出版发行
(北京和平里)
地质出版社印刷厂印刷
(北京海淀区学院路29号)
新华书店总店科技发行所经销

开本：850×1168 1/32 印张：12.4375 字数：328000
1989年10月北京第一版·1989年10月北京第一次印刷
印数：1—1570册 国内定价：4.65元
ISBN 7-116-00484-X/P·409

译 者 的 话

地质统计学是数学地质的重要分支。它首先由D·G·克立格工程师在南非的金属矿产储量计算中使用，后由法国马特隆教授领导的小组对此作了深入研究并系统地总结出地质统计学的理论和方法。地质统计学首先在矿石品位和矿产储量计算上引进了数理统计的方法，提出了变差图和区域化变量的概念，从而提高了矿产储量评价的精度，后来这种方法扩展到了地质学科的其它领域。

我国近几年来开始引进这种方法，正处在试验、研究和探索阶段，译出戴维教授这本书，以适应国内探讨地质统计学方法的需要。

本书主要介绍实际计算方法，对必要的理论只做了简练的介绍。书中列举了大量的矿山实例，并将一些实例分成了片段穿插于全书来讨论。

本书特别适合于矿山地质工作者和数学地质工作者阅读，也可做为地质学院师生的参考书。

本书的符号表及第一至第十章由孙惠文译出，其余部分由刘承祚译出。限于译者的水平，文中不妥之处，敬请读者批评指正，译者深表感谢。

目　录

符 号 表

大家知道，长期以来地质统计学家似乎具有这样一种能力，即能够判别出现在同一页中的相同符号所代表的不同含意。写这本书的时候，注意到了这种情况，并尽量在全书中保持符号的一致性。但还是存在一些问题，它们不可能有完美的标准符号，例如 γ是对数正态分布中中位数的标准记号，但它也是变差图的标准记号，类似的还有：α 一般用于表示对数正态总体中的对数平均值，但它也是德威依斯变差图方程中 $\gamma(h)=3\alpha\ln(h)$的标准记号。一个随机变量的累积概率常常记做 $F(x)$，但在本书中它也表示一个样品的方差！西舍尔t-估计因子已经使用了30年；我们也还有学生氏 t 分布和边界以上吨位曲线 $t(x_i)$。"a"可以表示斜率、变程、权系数或累积量，并且有时人们在同一报告中要使用所有这些量；因此，某些局部变化显然是必要的。然而，所有这些变量都使用不同的符号也不见得就方便，只要能做出肯定的判断，能把它们区别开来并能说出哪个符号是什么含意就可以了。

本书在讨论西舍尔，吉，英加梅勒斯等人的贡献时，仍然保留着他们的习惯符号，在每个特定部分给出了定义。

关于单位系统，本书保留了原有习惯，所选择的例子都用当初计算时所用的实际单位。下面以字母表的顺序列出所使用的符号及其所表示的意义。

$*$	上角标，指估计量
a	球型变差图的各向同性变程
$a_{标}$	在标明方向上球形变差图的变程
a_i	在一个线性组合中的权系数
a_{il}	在d_l^*的估计量中 $Z(X_i)$ 的第 l 个系数
A	$C(h)$ 的平均值（块金效应），或由上下文

	确定
〔A〕	克立格方程组中未知系数的向量
A_i	指标为 i 的样品的影响域
α	内蕴离散系数（3α＝德威依斯变差图的斜率）
α	在对数正态总体中对数值的平均值
B	块段
β	在对数正态分布中对数值的标准差
C_0	块金效应
$C+C_0$	基台
$C(h)$	随机函数的协方差
$COV(X,Y)$	X和Y的协方差
γ	对数正态总体的中位数
$\gamma(h)$	随机函数的变差图
$\gamma'(h)$	$\gamma(h)$的局部值
$\gamma_V(h)$	大小为V的样品的变差图
$\gamma(\theta,h)$	θ方向上的变差图
$\gamma_{ij}=\gamma(h_{ij})$	
$\overline{\gamma}(v,V)=\overline{\gamma}_{vV}$	v 和 V 之间变差图的平均值（h 的一端在 v 里，另一端在 V）
$\gamma_{同}(h)$	变差图的各向同性分量
$\gamma_{带}(h)$	变差图的带状分量
$\gamma_{RC}(h)$	变量 R 和 C 的交叉变差图
$\gamma_{Ml}(h)$	长度为 l 垂直于 h 的样品的变差图
d,D	由上下文所确定的维数
d_l	漂移的第 l 个系数
D_Δ	Δ 方向上直径变差
〔D〕	克立格方程组中右边的向量
δ	物质密度
Δ	方向

$\Delta h, \Delta\theta$	h, θ的增量
$E(X)$	X的数学期望值
$E(Y\|X)$	已知X时出现Y的条件期望值
$E(Y\|X=x_0)$	已知$X=x_0$时出现Y的条件期望值
ε_i	指标为i的一组估计误差
$f(X)$	随机变量X的概率密度
$f_l(x)$	漂移表达式中第l个函数
$F(X)$	随机变量X的累积分布
F（组）	组内点样品方差（组$=l, S, V, \cdots$）
$G(Z)$	高斯积分（从Z到∞）
h, H	由上下文确定的距离
$\mathbf{h}$	长度为h的向量
$h_{标}$	在标定方向上的h分量
h'	h对各向异性所作的修正
$h'_{标}$	在标定方向上h'的分量
h_{ij}	X_i点和X_j点之间的距离
i, j, I, J	由上下文确定的整数
k	由上下文确定
$k(x)$	一组特征函数
k'	修正因子
K_1, K_2	拉斯基常数
$K(\mathbf{h})$	随机函数的共变差图
l, L	由上下文所确定的长度
λ	克立格系数
m	总体的平均值
m_0	零边界矿床的平均品位
$m(x_c)$	边界以上吨位x_c的平均品位
$m(x)$	在点x（漂移）的期望品位
m_d	面内两个点之间的平均距离
μ, μ_i	拉格朗日乘数

n	样品数
N	负二项分布的参数
$N(h)$	成对样品数，成对点之间距离为 h
0	一个点样品
ω	正态分布截断后平均值计算中的因子
p	由上下文确定
$p(x)$	离散概率
P	负二项分布的参数
$P\{X\}$	X的概率
$P\{X\|Y\}$	已知Y时X的条件概率
$P\{X,Y\}$	X和Y的联合概率
$P_V(h)$	组V的几何共变差图
$q(x_c)$	x_c以上金属量
Q	负二项分布的参数
$Q(x_c)$	x_c以上金属量的比例
R^3	三维空间
R^n	n维空间
r	有限子样的相关系数
ρ	总体的相关系数
s,S	由上下文确定的面
s	一组样品的标准差
s_m	平均值的标准误差
s^2	一组样品的方差
$s^2(0/i)$	指标为 i 的块段内点样品的局部方差
σ^2	总体方差
σ_e^2	估计方差
σ_m^2	平均值的方差
$\sigma(X)=\sigma_X$	X 的标准差
$\sigma_{X_iX_j}=\sigma_{ij}$	X_i点和X_j点的品位协方差
σ_{VX_i}	样品X_i和集合V的品位协方差

σ_V^2	大小为V的块段的品位方差
$\sigma_B^2=\sigma^2(B/D)$	D矿床中块段B的方差
$\sigma(V_1,V_2/D)$	D内V_1和V_2集合的品位协方差
σ_N^2	由块金效应引起的方差
$\sigma_0^2=\sigma^2(0/D)$	矿床内点样品的品位方差
σ_{NV}^2	从N个样品平均值中估计块段V品位的估计方差
σ_s^2	曲面s的估计方差
$\sigma_{附}^2$	边缘效应的方差
σ_K^2	克立格方差
$[\Sigma]$	克立格方程组中σ_{ij}矩阵
t	西舍尔t-估计因子
$t_{(1-\alpha,N)}$	在N个样品和α风险下学生氏-t值
$t(x_c)$	边界以上品位x_c的总吨位
T_0	零边界矿床的总吨位
$T(x_c)$	边界以上x_c的吨位比例
T	时间区间
θ	方向
θ	三维对数正态分布中位置参数
θ	泊松分布参数
$\boldsymbol{u}$	单位向量
U	变差图处于准稳定的范围
v,V	由上下文确定的体积
$\underline{V}$	V中的一组体积
$\mathrm{VAR}(X)$	X的方差
w,W	子样的权值
X,Y,Z	随机变量
x,y,z	随机变量的值
$\boldsymbol{x}_i,\boldsymbol{X}_i$	空间的一个点
$\bar{x}$	一组样品的平均值

x_c	边界品位
$x_{ij}=x_i-x_j$	
$z(X_i)=Z$	点X_i的品位
Z_i^*	点X_i的估计品位
$Z(V)$	体积V的品位
$Z_c(X)$	点X的化学品位
$Z_R(X)$	点X的辐射度品位
$Z_s(X_\alpha)$	在点X_α的模拟品位
ψ,ϕ	旋转角

第一章　基础统计理论和应用

第一章和第二章内容提要　本书的前两章简要地介绍了统计学理论的基本方法，以及它们在矿产资源评价中的应用。

第一章的内容对于从未接触过统计学的读者了解后面讨论所需的基础知识已经足够了，而对专门从事数理统计工作的读者，我们只能表示歉意。作者认为，第二章所写的分布理论对矿产储量问题可能有帮助，但也可能写得不成功。已经熟悉统计学的读者，至少要阅读第二章，以便有把握正确地将统计学和矿业问题联系起来。

1.1　矿产储量计算中使用到的统计学术语

任何统计学教科书都是从一些基本定义开始，而人们往往容易忽视它们，事实上记住这些定义是非常重要的，这样能避免无意义的陈述。为了使这些陈述写得合适，我们先来复习一下这些定义，看看它们对我们的这一类的课题是多么重要。

1.1.1　全域

在一般的统计学教科书中没有这个词的定义；尽管这个名词还没有得到普遍承认，但在定量地质科学中我们需要它，并且会由此立即引出许多非标准的统计学问题。

全域是我们所涉及的物质的总体，同时也是所有可能数据的总体。全域可用一个属性或多个属性来说明其特征——它可以是一维的，也可以是多维的。在我们这里，比较容易确切地限定我们所涉及的范围——矿床。矿床可能有地质边界，这时，全域就是顶板和底板之间的物质；但到什么深度？全域可以将矿床定义

为自地表向下1000ft❶或5000ft，这是两个不同的全域。当仅有以品位圈定的边界，核心较富，其周围有浸染矿化时，问题就困难得多。这时，在对其他概念下定义之前，全域在空间中的位置是不清楚的。

1.1.2 抽样单元和总体

抽样单元是所度量的全域的一部分，它可以是一块5pound❷重的手标本，一块10ft长的岩心，或者一卡车矿石。当我们叙述全域的特征时，必须指出它的抽样单元。在日常生活的科学问题中，抽样单元通常是个体，例如：一个测量了体重的人，或一只老鼠等，这些是“天然的抽样单元”。在矿床中可能不是这样，同一个总的范围，即所谓全域，可以分成不同类型的抽样单元，这样就产生不同的总体，10ft样品的总体与100ft块段的总体可以

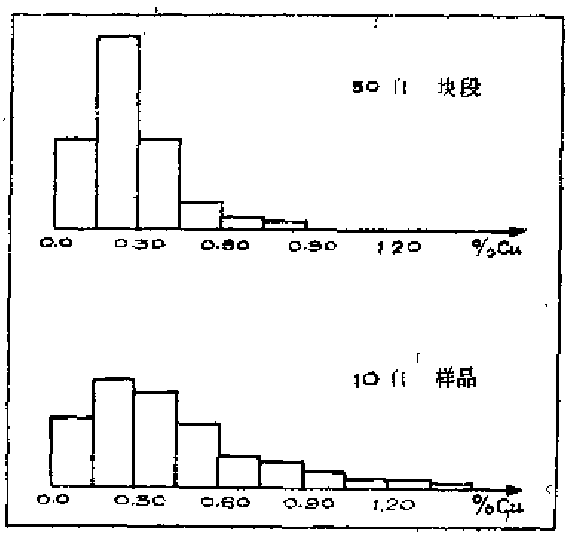

图 1 在斑岩铜矿床中10ft样品品位的直方图同50ft块段直方图作比较

---

❶ 1ft = 0.3048m；1in = 0.0254m

❷ 1pound = 0.4536kg

有很大的差别。我们懂得了如何表征总体的特点以后，这是不难理解的（图1给出了斑岩铜矿的一个例子）。

换句话说，抽样单元的大小是非常重要的，在没有指出品位单元时，不要去论述品位总体。

记住这样三个定义以后，现在我们这样看待统计学的问题：给定一个总体，在一系列抽样单元中，对一个属性的有限次数的测量，构成对这个总体的描述。在某些情况下，显然测量是在一个总体的样品中进行的，而结论希望是对另一个总体做出的，因此，人们必然要对这些结果产生很大的怀疑。

1.1.3 总体的特征

给定一个总体和一系列样品，人们总希望以简单的方法来表示一系列的测量结果，频率分布图正是起这个作用。给定一个属性区间，把落在这个区间的样品数做成简单的表格，然后，将这些数分成等级并画出图来。可以画出这样一种图形：将水平轴分成区间，在每个区间上画一个长方形，长方形的面积与落在此区间内的样品数目成正比，这种图形叫直方图，它代表数据的分布。因此，直方图是第一个手段。应用这一手段可以表示出在同一全域中块段品位的总体与样品品位的总体是不相同的，由图1看出，这还是相当明显的。

1.2 几点理论问题

1.2.1 随机变量

随机变量是这样一个变量，它可以取几种可能值中的一个值。当从矿床中抽取一个样品时，人们不知道这个样品的品位是多少，但是所有可能的品位是已知的。如果矿床中样品总数和具有已知性质（如：它的品位是在1.5%Cu和1.59% Cu之间）的样品数为已知，那么可以确定这个事件的概率。这是事件 A的经验概率定

义，它等于事件A可能发生的基本结果数与总的基本结果数之比，假若所有基本结果（所有样品）发生（取到）的可能性相等。

现在，我们不是仅仅考虑一个事件（一个品位区间），而是考虑这个事件（样品）所有可能的取值（品位值）。

1.2.2 概率分布

一次随机抽样的可能结果，可以用它的品位（随机变量）概率分布来描述。这个分布可能是已知的也可能是未知的。例如，投掷一个骰子，其可能结果是1、2、3、4、5、6，或者说6个数都以相等的概率1/6出现。在一个矿床中，人们希望知道在2—4%或6—8%，……，区间中得到一个品位的机会有多少。实际上，那样的分布从来不会知道。能够做的是计算经验概率分布，然后，设法从经验抽样分布中推断可能具有的理论分布。

假定变量只取整数值时，此时变量为离散的，分布将把每一个可能的值x同概率$p(x)$联系起来。显然，$p(x)$是非负的，所有$p(x)$之和等于1。在连续分布情况，对每个x，概率密度$f(x)$同x和$x+dx$之间的一个概率值$f(x)dx$有联系，其中dx是无穷小量。取值在a,b之间的x的概率是

$$\mathrm{Prob}\{a\leqslant x\leqslant b\}=\int_a^b f(x)\,dx$$

这个量也是非负的，如果积分区间是从$-\infty$到$+\infty$，这个量是$\int_{-\infty}^{\infty} f(x)\,dx=1$。把小于等于某个已知值$x_0$的概率叫做累积概率$F(x_0)$：

$$\mathrm{Prob}\{x\leqslant x_0\}=\int_{-\infty}^{x_0} f(x)\,dx=F(x_0) \tag{1.1}$$

下面结果成立：

$$F(-\infty)=0,\quad F(+\infty)=1$$