地质统计学

译文集

桂林冶金地质研究所情报室编

地质与勘探编辑部

地 质 统 计 学 译 文 集

桂林冶金地质研究所情报室编

地质与勘探编缉部出版
桂林市第二印刷厂印刷

787×1092毫米 32开本 3.5印张
字数140000 印数7500 1977年12月

说　　明

用算术平均法和常用的加权平均法计算矿体厚度、平均品位和金属储量时，由于未考虑样品的空间位置，计算结果往往与实际情况有出入。其中比较典型的一种误差是：对高品位块段的评价容易偏高，对低品位块段的评价则容易偏低。从事矿床勘探和开采的一些人由这个问题开始，逐步发展了数学地质学的一个分支，即“地质统计学”。它的基本特点，就是在资源评价和储量计算中用比较严谨的数学方法来处理在空间上展开的地质现象，包括定位和定向的样品的品位和矿体厚度参数。在计算技术高度发展的今天，用这种方法来处理勘探成果的大量资料和进行资源评价，已经不再有多大困难。因此，关键的问题就在于掌握地质统计学的原理及其具体方法。

为了便于地质人员在较短的时间内能对在最近二、三十年内发展起来的地质统计学有一个大致的了解，我们在这本小册子里，既收集了关于地质统计学的几篇早期论文，也收集了反映将地质统计学用于勘探实际的较新报导，最后还收集了一篇附有在电子计算机上应用克瑞吉法的程序的文章。

为了阅读这些文章，读者需要有常用统计学和微积分的基础知识。此外，如果能大体按本书编排的顺序逐篇分析这些文章，一定能按地质统计学本身的逻辑较快地掌握它。

目　录

地 质 统 计 学 原 理

G.Matheron〔法〕

就一般意义上说，地质统计学是研究对采矿工程师和地质专业人员有用的一些变量的空间分布规律。这些变量可以是品位、厚度、储量等等，它们对矿床评价问题有极其重要的实际意义。

从历史来说，地质统计学和采矿学是同时出现的。且开发者想要预测他下一步工作的结果时，尤其是开始取样、分析样品、按相应的厚度和影响范围计算加权平均品位时，人们就可以认为地质统计学诞生了。就反映矿化的空间特征来说，那些传统的方法仍然保持着它们的全部价值；现代的理论成就决不是要否定这些方法，而是要把它们作为出发点，并把它们提到更高的科学水平。

传统的方法虽然能够正确地求出平均值，但是却不能反映矿化强度的变异性（或离散度）这个重要的特征。几十年前，为了反映这个特征，曾经使用过经典概率论方法。尽管有时因机械地搬用这些方法而产生了错误的结果，但总地来说，结果还是有益的。就经典统计学方法不能涉及空间分布这一点来说，它实际上是有局限性的。事实上，南非研究所已发表的Krige和Sichel的成果中，常常说到他们采用了经典统计学，但他们的方法越来越不同于经典统计学，而是针对一定的目的进行了合理的调整的

统计学。

当明确地认识到经典概率论的不足和考虑矿化现象空间分布特征的必要性时，出现了第二次决定性的变化：实现传统方法和统计学方法的高度结合。因此，地质统计学就着手改造自己和使之数学公式化。这不是别的什么，而正是把现有的开采经验抽象出来，使之公式化和系统化。这一工作仍沿袭了统计学的术语（如方差和协方差），但这些概念已包含了新的内容。这一语汇上的相似性不应混淆其含义上的差别。经过长时间的发展，地质统计学理论已能处理在空间上展开的自然现象，而不只是随机事件。因此，它的方法与数学物理学方法（特别是调和 分 析 方法）很相似。

经典统计学概念的缺欠

简而言之，我们在下面虽只限于讨论矿体的品位，但是所得出的结论却具有普遍意义，并且对于在空间上展开的任何特征都适用。按常用的统计方法，在一个矿体上所采样品的品位可用直方图分级，这种方法不考虑样品在矿体里的位置。但是，只知道一个矿体中某一品位的频数是不够的，还必须知道不同的品位在矿区是怎样分布的，特别应该知道有经济价值的矿块的大小和位置。在文章的开头。我们提到这样一个事实：常用统计学不能考虑自然现象的空间特征，而这些特征确实是非常重要的。

严格地说，经典概率论研究的是博奕变量。纯粹的掷钱币出现正面或反面的例子清楚地表明经典统计学只能做些什么。让我们在每次钱币按正面落下时记＋1，相反时记－1。在掷出钱币前无法预测出现＋1还是－1；我们

只知道将要发生的是这两种可能之一，即非此即彼。一个博奕变量有两个基本性质：①可以无限次重复。对变量赋予数值的试验，至少从理论上说存在这个可能性。例如，我们可以按自己的需要掷钱币。②相对于前、后两次试验来说，每次试验都是独立的。如果前一百次试验都出现反面，第一百零一次试验出现正面的可能性仍为二分之一。

显然一个矿体中某一样品品位不具备这种性质。矿体的品位值首先是唯一的。只要矿块被开采了，就不可能再次重复。当要求了解一个样品的品位时（例如这是一个规格为已知的槽样），由于具有（x,y）坐标的样品的品位是唯一确定的，所以，结果也是唯一的。但是，可以在紧靠第一个样品的一旁采第二个样、第三个样等等……，这显然说明有重复试验的可能性。实际上，这并不是严格相同、而是稍有差别的试验。即使有重复出现的可能，第二个性质也肯定不具备：两个相邻的样品不一定独立。平均地说，如果这两个样品是从一个高品位矿块上采取的，两者将都有达到高品位的趋势，反之亦然。这种不同程度的强化趋势表示在矿化空间内品位变化具有程度不同的连续性。

如不理解上述事实和粗枝大叶地搬用经典统计学，有时会导致惊人的谬误。近五十年来，在资源勘探工作中，经验提醒人们要绘制每个钻孔的位置图（即准确地标定它们的位置）。采矿人员仍继续按传统的规则网格采样，地质统计学后来也证明这是正确的。如若不然，就有认为一个矿体评价工作的精确度只取决于样品的数目（与他们的位置无关），而且随着样品数的平方根变化。生硬地套用误差理论，会产生荒谬的结果，例如，如果用钻孔揭露某

一矿体，要多采100倍的样品才能使精度提高10倍；而用长度为5毫米的岩心代替长度为50厘米的岩心就足够了，这当然是错误的。用这种办法所得到的样品品位多半是虚假的；除了无限重复相同的信息外，并不能得到什么别的东西。地质统计学说明，5毫米长的岩心和50厘米长的岩心的品位资料精度是相同的，每一个采矿人员自然能理解这一点。

区域化变量的概念

因此，品位无论如何不能与博奕变量等同起来。为了强调自然现象的空间特性，我们将明确地提出区域化变量。区域化变量是一个实函数。在空间的每一点上，它都有确定的值。

在一般情况下，这种函数的性质太复杂，不易用一般的数学分析方法来研究。从物理学或地质学的观点看来，某些定性特征是与区域化变量的概念相关连的。

（1）首先，一个区域化变量是有局限性的。它的变化只限于矿化空间之内（矿体或矿层的范围之内），这一矿化空间称为区域化的几何域。况且，这个变量一般是按几何支架定义的。在只考虑矿石品位的情况下，这个支架不是别的，而是具有一定的几何形态、规格和方位的样品的体积。如果在同一个矿体里改变这一几何支架，即将得到另一个区域化变量。它虽与第一个变量类似，但并不与它们吻合。

换句话说，10公斤岩心样与10吨破碎的矿样是不同的。同时，也常常需要考虑到点支架的具体情况。例如，采样点落到无矿处还是矿化颗粒上，点品位值将分别取0

或1。

（2）其次，这个变量反映出它在空间上的变化有不同程度的连续性，这可通过两个相邻样品之间的变差来表示。赋予某些具有几何特征（地质体的厚度或倾向）的变量以严格的数学连续性。但这种变量（品位或储量）往往只具有较松驰的连续性。换句话说，这只是平均意义上的连续性。在某些情况下，连这个平均连续性也不成立，因此我们应提到跃迁效应。

（3）最后，这个变量能反映出不同类型的各向异性。可能存在一个特定的方向，沿着这个方向品位没有明显的变化，而沿着其他方向品位将迅速变化。这种现象称为分带性。

对任何区域化变量而言，特殊的变异性可以迭加在一般的规律上。例如，在沉积矿床中应注意到分层效应。厚大的沉积层内可能包含若干彼此可分的含矿层。在每个含矿层内，矿层沿垂直于层位的方向依次排列并被不连续的界面所分隔。在一个特定的层位中，品位几乎是不变的。而从一个层位过渡到另一个层位，品位将突然变化。这一常见的众所周知的现象，看来好象是一个理论的公式化的问题，但仍然是基本的问题，不考虑它就抓不住要领。由解理加剧了的垂向上的不连续性与横向上的不连续性也会同时出现。这种横向上的不连续性存在于透镜状矿层的边缘。如果有矿层迭加的现象，它将反映出在每个层位上沉积区被分为一些天然的小矿集。在计算过程中，通过品位的极限效应即可发现它。

在网脉型矿床里，同样能见到高品位矿脉或浸染状高

品位团块，这个网脉效应正如分层效应和矿层迭加效应一样，表示为在均匀的几何区间存在不连续的网格。从回然不同的角度来说，跃迁效应也是同一性质的现象，在这里，不连续的网格是矿块间的夹石。

区域化变量的这些特殊性应由地质统计学来通盘研究，因为经典概率论是无能为力的。由于采用了变程方差函数这种简单的数学工具，这一点是可能办到的。

变程方差函数

变程方差函数可以表征矿化的连续程度。按经验方法，在横坐标上取距离d，纵坐标表示彼此相距为d的样品品位差的平方的平均值。令f(M)为区域化变量在几何域V中M点的值，此区域化变量是用特定的几何支架v（一般来说，这个几何支架v很小，它的极限可以看成一个点）定义的。由向量h定义的半变程方差函数γ(h)或离散律为：

$$\gamma(h)=\frac{1}{2V}\iiint_{v}[f(M+h)-f(M)]^{2}dV \quad \cdots\cdots ①$$

变程方差函数是随距离h增大而增大的函数，因为两个样品相距越远，它们的品位差也越大。它赋予传统的关于样品影响域的概念以确切的解释。变程方差函数值程度不同的突然增长，确切地表明了某已知样品对距离越来越远的盘区的影响突然变化的程度。变程方差函数很好地反

映了区域化的定性特征：

（1）γ(h)在原点附近的稳定性，表现了矿化强度的变异性，它大致可分为四类（图1）。

第一类，变程方差曲线在原点附近趋向于抛物线时，代表诸如矿层厚度这样具有高度连续性的区域化变量。

第二类，或线性型，其特点是在原点附近有斜向的切

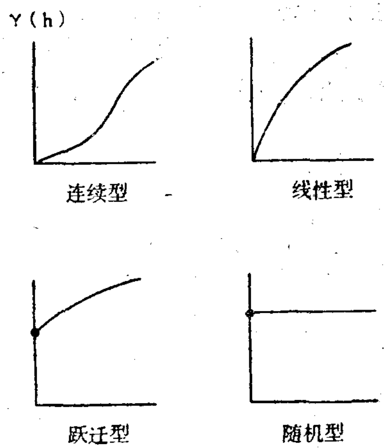

图1

线，它代表变量有平均的连续性。在金属矿床中常有这种类型的品位变化。

第三类，变异曲线在原点不连续，这表明变量甚至在平均的意义上也是不连续的，而是跃迁性的。

第四类是相当于经典的随机性变量的极端情况。在第一类（连续型函数）和第四类（纯随机型函数）间有一个

过渡区，对这一过渡区进行研究，也就是地质统计学的真正课题。

（2）沿着空间上的不同方向变程方差函数是不同的。按①式定义的函数$\gamma(h)$不仅只取决于向量h的长度，而且还取决于它的方向。通过研究不同方向上变程方差曲线的差异性，可以揭示优异方向、富矿脉或富矿柱。

（3）构造特征也能在变程方差曲线上反映出来，例如，矿层交替现象在实验曲线上将表现为变程方差曲线在超过一定距离后变为水平延伸线，这段距离大约等于沉积

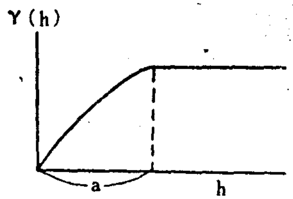

图 2

形成的矿巢的平均直径。在不同的方向上这种距离不同，因而有可能确定富矿巢的长轴方向和大致形状。

变程方差曲线既不表示矿化现象总体，也不表示它的细节，它只是反映它们的主要特征的综合形式。振动现象的简谐分析并不给定每个谐波的相位和振幅。振动的局部外形主要由相位决定，而能量只与振幅的平方有关。振幅谱曲线不描述整个现象，而只说明主要的即能量的特征。变程方差曲线（更确切地说，是它的富里叶变换曲线）起着与这个振幅谱曲线完全同等的作用。

在下节，列出了变程方差曲线的几种可能的用途。对它的系统研究显然超出了本文的范围。作者只能提出一些例子和几个主要公式。更详细的内容请读者参考《应用统计学概论》。

绝对弥散律或本征离散律

按①式定义的半变程方差函数只在区域化变量的几何域V上成立。如果不考虑整个V，而是只考虑它的一部V′，可能得到一个与γ(h)不同的γ′(h)函数。但是，我们有一个直观的想法，即在按地质观点认为是均匀的几何域内，可能存在一些固有的、不依存于位置的、表示区域化变量变异性的因素。

如用公式来准确地表示，那个直观的想法将导出一个绝对弥散律或本征离散律的假定，它可表示如下：

$$\gamma'(h)=\gamma(h)$$

这意味着变程方差函数不取决于为计算它而在矿床V中选定的V′。立刻可以肯定，这个假定对这一理论的发展并不重要，而且由于某些数学的复杂性可能将它抵消。

虽然如此，这个假定仍能使理论叙述大为简化，因此在这里还将沿用它。变程方差函数在空间上的缓缓偏移一般可通过经验确定。如果这个偏移不太明显，用绝对弥散律假定所得到的结果将与实际情况非常近似（当实际采用的γ(h)是从矿体的具体部位算出来时就是这样）。

当这个假定成立时，半变程方差函数γ(h)本身将获得一个固有的意义，并常常称之为本征（或绝对）弥散律。简而言之，它就是区域化变量的本征函数。

方 差 和 协 方 差

首先让我们设想有一个区域化变量（比如说品位），它是按点支架定义在区域V上的，并且服从于本征弥散律$\gamma(h)$。令$f(M)$表示在V上M点的品位值。我们通常不研究一个点的品位$f(M)$，而关心在M点（v的重心位于M点）所取的已知其大小、形状和方位的一个样品v和品位$y(M)$。这个新变量是通过在以M为中心的体积v内对上述的点品位积分而推导出来的。

$$y(M)=\frac{1}{V}\int_V f(M+h)dv \quad \cdots\cdots\cdots\cdots\cdots ②$$

必须用这个变量来建立一个测量$y(M)$在V内的弥散度的参数。与经典概率论一样，这个参数可称之为方差。在V内，点品位的均值为m，

$$m=\frac{1}{V}\int_V f(M)dV$$

在v内，$y(M)$的方差定义为$[y(M)-m]^2$在V内的均值，令

$$\sigma^2=\frac{1}{V}\int_V [y(M)-m]^2 dV \quad \cdots\cdots\cdots\cdots\cdots ③$$

应当指出，这个概念在最初只有几何意义而没有概率意义。但这并不妨碍我们在应用过程中采用常规统计方法从实验数据中计算这些方差。象在积分式③中一样，如果按空间顺序来求σ^2，或者预先重排直方图，则相同的表达式$(y-m)$在二次计算过程中应乘上相同的权。但是，根据基本概念，定义③还具有统计量所没有的物理意义。根据变程方差函数的表达式①、②和方差的定义③，交换一下积分次序后可得：

$$\sigma^2=\frac{1}{V^2}\int_V dV\int_V \gamma(h)dV'-\frac{1}{v^2}\int_v dv\int_v \gamma(h)dv' \quad \cdots\cdots ④$$

每一个六重积分都有非常确定的意义：它代表当向量h的两端扫过体积V(或v)时，γ(h)在V(或v)内的均值。

如果我们写出：

$$F(V)=\frac{1}{V^2}\int_V dV\int_V \gamma(h)dV,$$

即F(V)＝γ(h)在V内的均值，即可得出：

$$\sigma^2=F(V)-F(v) \quad \cdots\cdots ⑤$$

于是，根据点品位变程方差曲线的概念可以预先计算在矿床的任何部分V内任何样品v的方差。应当指出，这个方差不仅取决于体积v和V的大小，而且与它们的形状和方位有关。

关系式④的物理意义是很有启发性的。一个宏观的样品v（可以视作由大量微小样品dv累积而成）的方差完全不决定于微小样品的个数，也不取决于它的方差，而仅仅与本征函数γ(h)在几何域V中的平均值有关。按经典统计学（把这些微小样品看成独立的）求出的方差应为1/v，实际上不存在这样一个矿床：由这个矿床取出的10吨爆破样的方差比10公斤岩心样的方差低一千倍。公式④说明了为什么会这样。微小样品的品位完全不是独立的，它们之间存在一个空间上的相关阵。而且，按平均意义说来，微小样品的品位与经典统计学的计算结果不同，因此，10吨爆破样的方差比10公斤岩心样的方差的千分之一要高得多。

可以看出⑤式的方差表达式具有可加性，如果我们考虑的是区域V中的矿块V′和样品v，并且指定$\sigma^2(V',V)$、$\sigma^2(v, V)$和$\sigma^2(v, V')$分别为在V内的V′的方差、在V内的v的方差和在V′内的v的方差，我们将得到：

$$\sigma^2(v, V)=\sigma^2(v, V')+\sigma^2(V', V)$$

这个公式称为克瑞吉公式。它是D.G.Krige在品位分布服从对数正态分布律的情况下导出的，它成立与否同特定的统计分布律无关，而只是与本征弥散律的存在有关。

除了方差之外，在地质统计学中还引入了协方差的概念，如果y(M)和z(M+h)表示以M点和M+h点为中心的样品V和V′的品位，y和z的协方差（在V内）作为h的函数定义如下：

$$\sigma_{yz}=\frac{1}{V^2}\int_V [y(M)-m][z(M+h)-m]dV$$

可以通过变程方差函数将它用类似于④的关系式表示：

$$\sigma_{yz}=F(V)-\frac{1}{vv'}\int_v dv\int_{v'}\gamma(k)dv' \qquad \cdots\cdots⑥$$

第二个积分表示当向量k的两个端点（一端与另一端的距离为h）分别扫过体积v和v′时$\gamma(k)$的平均值。作为一个特殊情况，我们来分析一下各向同性的代威依斯模型，并用下述各向同性的本征函数来定义它：

$$\gamma(r)=3\alpha\ln r \qquad \cdots\cdots⑦$$

其中$r=|h|$代表向量h的模，换句话说，它是M和M+h两点之间的距离。ln代表自然对数，参量α称为绝对弥散系数。它表明：品位的弥散性与样品和矿体的形状、体积无关。在样本v和矿体V同形的特殊情况下，由公式④和⑦可

得：

$$\sigma^2 = \alpha \ln \frac{V}{v} \quad \cdots\cdots\cdots\cdots\cdots\cdots ⑧$$

这就是代威依斯公式，它表示的是相似性原理。样本与矿体不同形时，一般不能用这个公式。但可以把任何几何形体v与它的直线型等价线d联系起来，直线等价线d可写为下列关系式

$$\ln d - \frac{3}{2} = \frac{1}{v^2} \int_v dv \int_v \ln r dv' \quad \cdots\cdots\cdots\cdots ⑨$$

公式④要求在任何矿体里样品v与长度为d的直线型样品有相同的方差，如果D和d分别为矿体和样品的直线型等价线，方差即可化为：

$$\sigma^2 = 3\alpha \ln \frac{D}{d}$$

对于某些几何形体的线型等价线已作了计算，并已作成了表。此外，我们还推出了一些简易的近似公式。例如，对边长为a和b的矩形有

$$d = a + b$$

对边长为a、b和面积为S的平行四边形有：

$$d = \sqrt{a^2 + b^2 + 2S}$$

对边长为a、b、c和面积为S的三角形有：

$$d = \sqrt{\frac{a^2 + b^2 + c^2}{3} + 2S}$$

如梯形上下两边为$a = \frac{L+l}{2}$和$b = \frac{L-l}{2}$，中线为m，面积为S时，