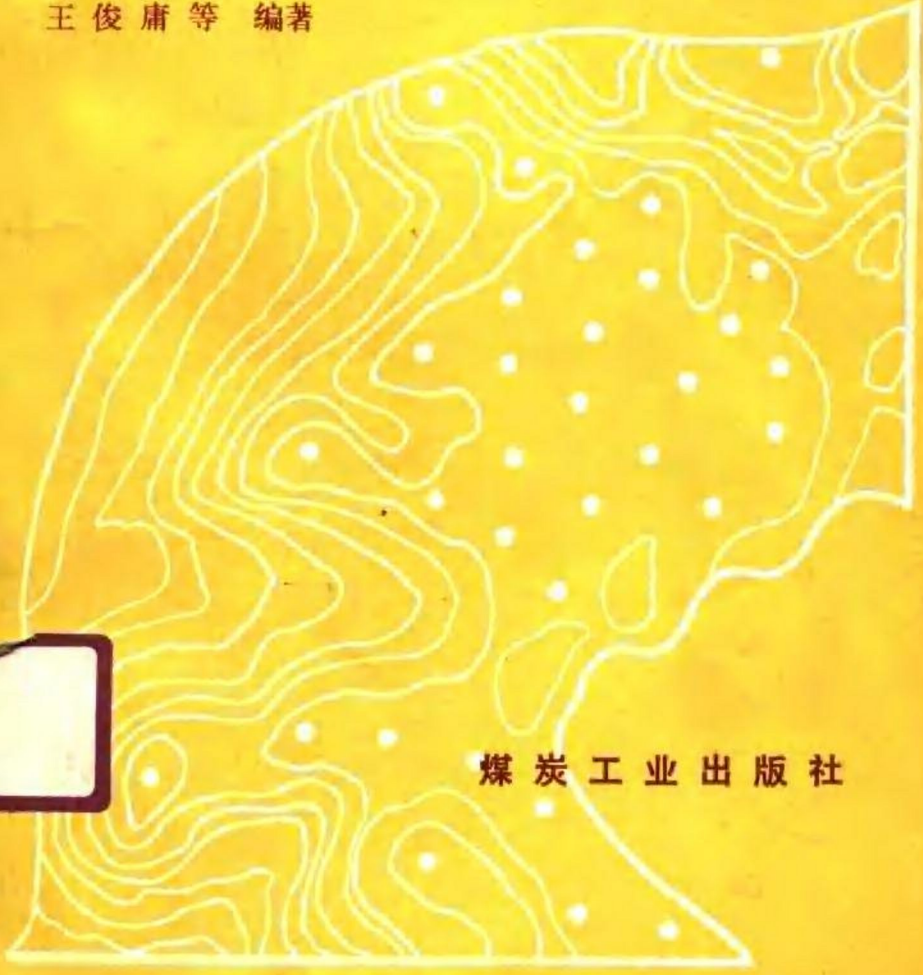


地质统计学及其在 煤炭资源开发中的应用

王俊庸等 编著



煤炭工业出版社

内 容 提 要

本书通俗地阐述了地质统计学的基本原理——线性地质统计学，并从实用角度结合煤田勘探和煤矿设计，介绍了地质统计学在煤炭资源开发中的应用。

全书分为两篇。第一篇为线性地质统计学基本原理，包括概论、区域化变量及其半变差函数、结构分析、普通克里格法、泛克里格法等；第二篇通过三个实例介绍了地质统计学在煤田勘探类型划分、勘探基本网度优选、煤矿床储量计算、煤质预计、储量分级以及勘探程度评价、建立煤矿床模型等方面的应用，对定量评价地质勘探精度进行了可行性探讨。

责任编辑：罗 醒 民

地质统计学及其在煤炭资源开发中的应用

王 俊 庸 等编著

煤炭工业出版社 出版

(北京安定门外和平里北街21号)

煤炭工业出版社印刷厂 印刷

新华书店北京发行所 发行

开本 787×1092mm¹/32 印张 10⁷/3 插页 4

字数 239千字 印数1—890

1990年11月第1版 1990年11月第1次印刷

ISBN 7-5020-0414-9/TD·375

书号 3204

定价 4.80元

前 言

地质统计学是60年代初期产生并发展起来的数学地质的独立分支，具有极强的实践性。从某种意义上说，它是一种专门研究区域化变量分布规律和估值的技术，在许多西方国家被誉为最优化的储量计算方法。它以其严谨的数理，定量化地描述了地质变量的变异特征，并给出精确的估值及估计方差，因此深受地质、矿业工作者的喜爱和推崇。目前，它的应用已经远远超出储量计算的范围，向着更广的领域渗透和发展。

在我国，地质统计学的传播始于70年代末。80年代初，煤炭系统也正式开始了地质统计学的系统学习和研究，直到现在，已经取得一批应用成果。为在煤炭系统推广和普及地质统计学，本书拟以煤矿床为对象，从实用角度深入浅出地阐述线性地质统计学的基本原理及其在煤炭资源开发中的应用。

本书的编写主要参考了王仁铎、胡光道的教材——《线性地质统计学》。

书中的第一篇由王俊庸执笔；第二篇的实例一由陈朝阳、雷宝林执笔，实例二由宋国强、王俊庸执笔，实例三由曹显煌执笔，并由王俊庸对全书进行统稿，第一篇和实例二的插图由宇桂珍清绘。曾澜审校了全部书稿并提出了许多有益的修改意见。

敬希读者对本书的错误和不当之处提出批评指正。

作 者

一九八七年十二月

目 录

第一篇 线性地质统计学基本原理

第一章 地质统计学概论	1
一、储量计算方法的演进与地质统计学的产生	1
二、地质统计学的理论基础和主要内容	18
三、线性地质统计学及其特点	21
四、地质统计学在我国的传播和应用	23
第二章 区域化变量及其半变差函数	25
一、随机变量与随机函数	25
二、区域化变量	32
三、半变差函数和协方差函数	36
四、估计方差	41
五、离散方差	48
六、正则化	52
第三章 结构分析	55
一、实验半变差函数	55
二、半变差函数的理论模型及其拟合	65
三、套合结构	76
第四章 普通克里格法	97
一、克里格方程组	97
二、克里格方差	103
三、对普通克里格法的几点认识	104
四、普通克里格法的应用技巧	113

五、整体估计问题	131
六、辅助函数	145
第五章 泛克里格法	160
一、泛克里格法概述	160
二、泛克里格法的半变差函数	166
三、泛克里格方程组	168
四、漂移 $m(x)$ 的估计	174
五、泛克里格法计算实例	177
第六章 地质统计学在煤炭资源开发中的应用	180
一、在煤矿床研究中应用地质统计学的特点	180
二、计算煤炭储量与估计煤质	183
三、优选煤田勘探基本网度, 合理布置勘探工程	198
四、储量与煤质的控制程度评价及储量分级	210
五、揭示地质变量在空间的变异规律	216
六、建立煤矿床模型	221

第二篇 应用实例

实例一 东荣矿区三井田的地质统计学研究	225
实例二 黑岱沟露天矿首采区煤田地质 资料的设计评价	276
实例三 地质统计学在河南煤田地质勘 探中的应用	303
主要参考文献	341

第 一 篇

线性地质统计学基本原理

地质统计学经历了孕育、产生和发展 3 个阶段的演进。30 多年来，已经积累了相当丰富的实践经验，当形成了比较完整的理论方法体系。作为一个新兴的工程学科，一个地质学与统计学相结合而产生的边缘学科，它的理论基础是区域化变量及其半变差函数。在其已有的几种研究方法中，线性地质统计学是最基本的也是应用最广泛的方法。本篇阐述线性地质统计学基本原理，即着重介绍区域化变量理论、半变差函数及普通克里格法。

第一章 地质统计学概论

一、储量计算方法的演进与地质统计学的产生

1. 储量及其计算方法的演进

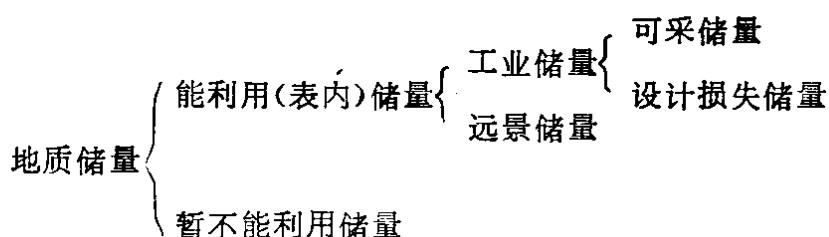
1) 储量的概念

储量，对于从事矿床勘探和矿业开发的人来说，显然是个非常熟悉的技术词汇。人们通常把储量理解为有用组分或矿石在地下埋藏的数量，并且认识到储量不仅是个数量的概念，而且还具有矿石质量、开采技术条件等方面研究程度的含意，因此它是勘探成果的集中表现。对于矿山而言，储量

则是矿山设计、建设与生产的基础，它的数量与质量直接关系到矿山的规模和生产发展的前景。总之，储量是矿床勘探的成果，矿山开发的基础和依据，对矿山具有极大的重要性。

在煤炭工业中，煤矿储量是指井田（矿田）内所埋藏的具有工业价值的煤炭数量。所谓具有工业价值主要指厚度大于最低可采厚度和灰分小于最大允许灰分含量的煤层。

根据煤炭资源能利用的程度和开采技术条件，对煤炭进行了如下的分类：



其中的工业储量是可以作为设计和投资依据的那部分能利用储量。它包括可采储量和设计损失。

在煤田勘探中，又根据对煤层勘探和研究程度的不同，以及设计、生产部门的需要将储量划分为 *A*、*B*、*C*、*D* 四级，其中 *A* 级和 *B* 级是高级储量。

从上述对煤炭储量的定义及分类、分级中可以看出：

① 传统的储量概念是将储量当作一种确定性变量定义的；

② 传统的储量概念仅能用储量分级定性地描述对矿床（煤层）勘探和研究的程度，没有储量精度的定量指标；

③ 传统的储量概念将矿石在地下实际埋藏的数量（真值）与计算得出的数量（估计值）未加区分。

2) 储量计算方法的演进

为了计算储量，人们先后提出和使用了多种储量计算方法。这些方法都是把精确地计算平均品位 ($\bar{c}$) 或平均厚度 ($\bar{m}$) 作为其主要目标。

最初人们是用算术平均法求平均品位或平均厚度的，在此基础上发展使用了地质块段法。算术平均法是把各取样点的观测值当作等权，而且彼此在空间上互不相关看待的。继而，人们又开始使用一种多边形法计算储量，它把多边形内的唯一观测值看作是該多边形的品位（或厚度）的平均值，实际上是一种最近点原则。后来人们发现空间各观测值是相关的，而且这种相关性的 size 主要由距离的远近所决定。待估点（或块）距取样点越近，受取样点的影响越大；反之，取样点对待估点（或块）的影响就小。基于上述认识，一系列的按距离幂次反比加权的储量计算方法应运而生。特别是电子计算机的应用，使这类方法得以迅速传播。它的基本计算公式为

$$C_B = \sum_{i=1}^n \left(\frac{c_i}{d_i^m} \right) / \sum_{i=1}^n \frac{1}{d_i^m} \quad (1-1-1)$$

式中 C_B ——待估块段 B 的平均品位（或平均厚度）；

c_i ——各取样点的品位值（或厚度值），取样点 $i = 1, 2, \dots, n$ ；

d_i ——各取样点与待估块段 B 之间的距离；

m ——上述距离 d 的幂次， $m \geq 1$ 。

由于距离幂次 m 的不同，则有不同的距离幂次反比法（图1-1-1）。习惯上最常使用的是 $m = 2$ 的距离平方反比法。图中有距离为100的两个观测点 A ， B ， A 的观测值为100， B 的观测值是0。当 m 由1变到 ∞ 时，两点 A 、 B 间变量值的分布情形如各曲线所示。其中3种极端的情形是：

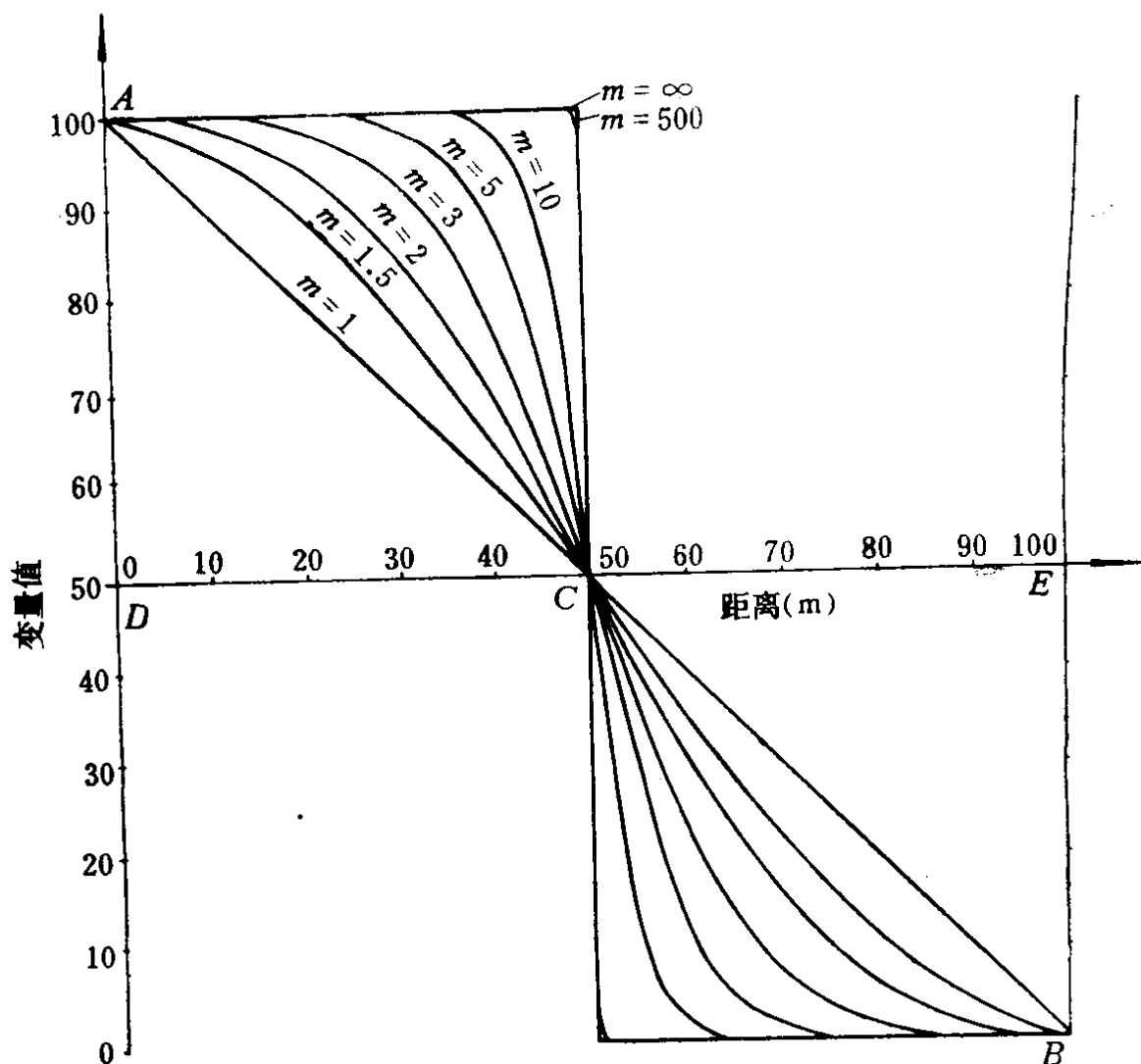


图 1-1-1 距离幂次反比法曲线图

$m = 0$ 时, A 、 B 两点间各处的内插值均相等, 为 A 加 B 的算术平均值 (直线 $D-C-E$), 即是算术平均法;

$m = \infty$ 时, A 、 B 两点间的内插值以中间点 C 为分界, 靠近 A 点的插值为 A 的变量值 (100); 靠近 B 点的插值为 B 点值 (0), 这即是多边形法;

$m = 1$ 时, A 、 B 两点间为线性插值 (直线 $A-C-B$)。

进一步的研究表明: 距离幂次反比法的幂次 m , 视计算变量的变异性强弱而变化, 变异性强的变量所取的 m 应该偏大; 变异性弱的变量所取的 m 就应小些, 一般 $1 \leq m \leq 3$ 。

通常惯用的是距离平方反比法 ($m = 2$)。可见, 在使用距离幂次反比加权法时, 有一个对幂次 m 的优选问题。怎样优选计算不同变量时所用的幂次呢? 1969年南非的牟克(MOC)提出了专门用于计算煤矿储量的加权滑动平均模型, 解决了幂次的优选问题。他所用的加权公式是

$$r_{x,y} = \frac{\sum_{i=1}^n \frac{r_i}{(a^m + d_i^m)}}{\sum_{i=1}^n \frac{1}{(a^m + d_i^m)}} \quad (1-1-2)$$

式中 $r_{x,y}$ ——点 (x,y) 处函数的估计值;

r_i ——已知信息点 (x_i, y_i) 处的函数值, $i = 1, 2, \dots$
 $\dots, n$;

n ——已知信息点的个数;

d_i ——已知点 (x_i, y_i) 和待估点 (x, y) 之间的距离;

a ——常数, 起光滑因子的作用, 是观测误差的大小和函数局部变化的函数;

m ——函数的区域稳定性指标, 距离 d_i 的幂次。

其中 d_i 、 r_i 和 n 是已知的。此模型的关键是能否给出 a 和 m 的优选值, 对于不同的煤层将有不同的 a 和 m 。

具体确定 a 、 m 的方法如下:

① 设在 n 个观测点测得煤厚的观测值, r_i 为 (x_i, y_i) 点的煤厚。在给定某个 a 、 m 值的条件下, 用除了 i 点以外的其它观测点 $(n-1)$ 的值以加权滑动平均模型(式1-1-2)计算 i 点的估计值;

② 按①的作法对所有的 n 个观测点都这样进行估值计算, 得到

$$P(a, m) = \sum_{i=1}^n [r_i - r_i(a, m)]^2 \quad (1-1-3)$$

式中 $P(a, m)$ ——在给定一组 a, m 值时, 用①、②步骤算出的各观测点估计误差的平方和;

$r_i(a, m)$ ——在给定一组 a, m 值时, 用①、②步骤算出的加权滑动平均估计值。

③ 多次改变 a, m 的大小, 重复①、②的计算, 得出多个 $P(a, m)$ 值, 从中选出最小的 $P(a, m)$, 此 $P_{\min}(a, m)$ 所对应的 a, m 值即是优选的常数光滑因子和距离幂次。用此经过优选的 a 与 m 代入加权滑动平均模型计算煤矿储量, 会得出令人满意的计算结果。

仔细分析式 1-1-2 可知, 此法的核心仍是距离幂次反比加权法。当 $a = 0$ 时 (即不考虑光滑因子的作用), 式 1-1-2 与式 1-1-1 基本相同; 若 $a = 0, m = 2$ 时, 即是距离平方反比法。因此, 可以将上述优选 a 和 m 的方法用于式 1-1-1, 进行距离幂次的优选。事实上, 已有不少人在储量计算时, 正式使用了这种优选距离幂次的方法。现在用地质统计法计算储量时, 通常用“交叉验证法”优选半变差函数的主要参数, 这种“交叉验证法”, 实质上就是优选 a, m 方法在地质统计学储量计算中的运用和发展 (后面将介绍“交叉验证法”)。

我们可将上述基于几何法的所有手工的和计算机实现的储量计算方法统称为传统储量计算方法。当然, 随着应用数学和电子计算机的推广和普及, 近年来又出现不少储量计算的新方法, 如趋势与偏差插值叠加法、双三次样条函数法等, 在此不拟赘述。

3) 传统储量计算方法存在的问题

传统储量计算方法各具特点，在长期的实践中曾被人们承认和使用，但它们仍存在着不足之处。具体表现在：

① 将计算储量的主要参数（如品位、煤厚等）看成确定性变量，忽视了这类变量所具有的随机性特征，只强调这类变量在空间上的相关性。于是仅按其间的距离所反映的相关性，构制出多种以距离加权的储量计算公式，这些计算公式均以确定性解析式为特点。实际上，决定储量的主要参数多具有两重性——既具有空间上的相关性又具有取值上的随机性。

② 传统储量计算方法仅考虑矿床的几何形态以及观测点和待估点(块)之间的距离因素，因之均属几何法。有的仅用局部观测点的信息值去推算待估点的可能值（如地质块段法、多边形法），忽视了这些参数——地质变量的变异规律性及其异向性，因而导致储量计算的精度不高。例如，煤层厚度在走向和倾向上经常有不同的变化规律，一般走向上的厚度变化小于倾向上的厚度变化，因此处于不同方位但与待估块 V 距离相等的4个已知观测点对待估块影响的权重就不会相同。图1-1-2中4个已知点的煤厚值是 Z_1 、 Z_2 、 Z_3 、 Z_4 ，待估块 V 位于其间而且与它们等距，但由于 Z_1 、 Z_3 处于待估块 V 的走向方向，变异性弱，而 Z_2 、 Z_4 处于 V 的倾向方向，变异性强，因而影响待估块 V 的权重 Z_1 、 Z_3 较 Z_2 、 Z_4 要大些。但是，用传统的距离幂次反比加权法计算时，将对 Z_1 、 Z_2 、 Z_3 、 Z_4 赋以相等的权重。

③ 传统储量计算方法不能给出储量计算精度的定量指标。从理论上说，储量计算方法的优劣应由计算精度决定。严格地说，对于没有计算精度定量指标的储量计算方法，是无法评价其优劣的，而不能给出估计精度的储量，也是没有

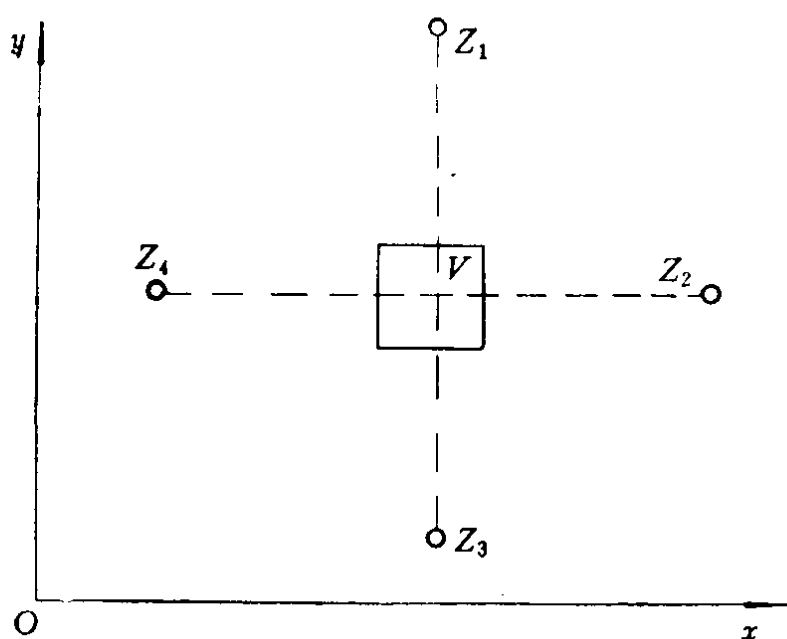


图 1-1-2 变异的异向性对信息权重的影响

使用价值的储量。但长期以来，由于计算技术和方法的局限，只能依靠经验、定性地去评价储量计算的精度及计算方法的好坏。为了检验储量的可靠性，通常借助于同种方法重复多次计算，或用不同方法校核来得到储量的相对精度。

④ 储量计算是对地下埋藏的矿石数量（真值）进行估计，储量是对这一真值的估计量。传统储量计算方法将真值和估计值混同一体，不加区分。因此，无法引进关于估计的科学方法（最小二乘估计、最大似然估计等），更无从使用估计误差、估计方差的概念。

⑤ 任何样品值都和取样的大小、形状有密切的关系，在地质统计学中，将样品或块段的大小与形状称为几何支集（Support），有的书上称为几何支撑或承载。通常的储量计算方法简单地把取自不同几何支集（以下简称支集）的样品观测值混同起来，不加区分地延伸到待估矿块上去，这

是极其不当的。因为一个钻孔的岩芯煤厚值和一个很大的待估块段的平均煤厚是不能等同看待的（后面将会讲到）。这一点对于复杂矿床的储量计算尤其重要。如南非的地质、采矿学家 D.G.克里格（D.G.Krige），在用多边形法计算南非某金矿的储量时发现：计算出的矿块品位与开采后的矿块实际品位之间存在有系统的偏差。即对富矿带的品位估计偏高；对贫矿带的品位估计偏低。这正是传统的多边形法用一个钻孔的平均品位代替多边形矿块的平均品位所致。

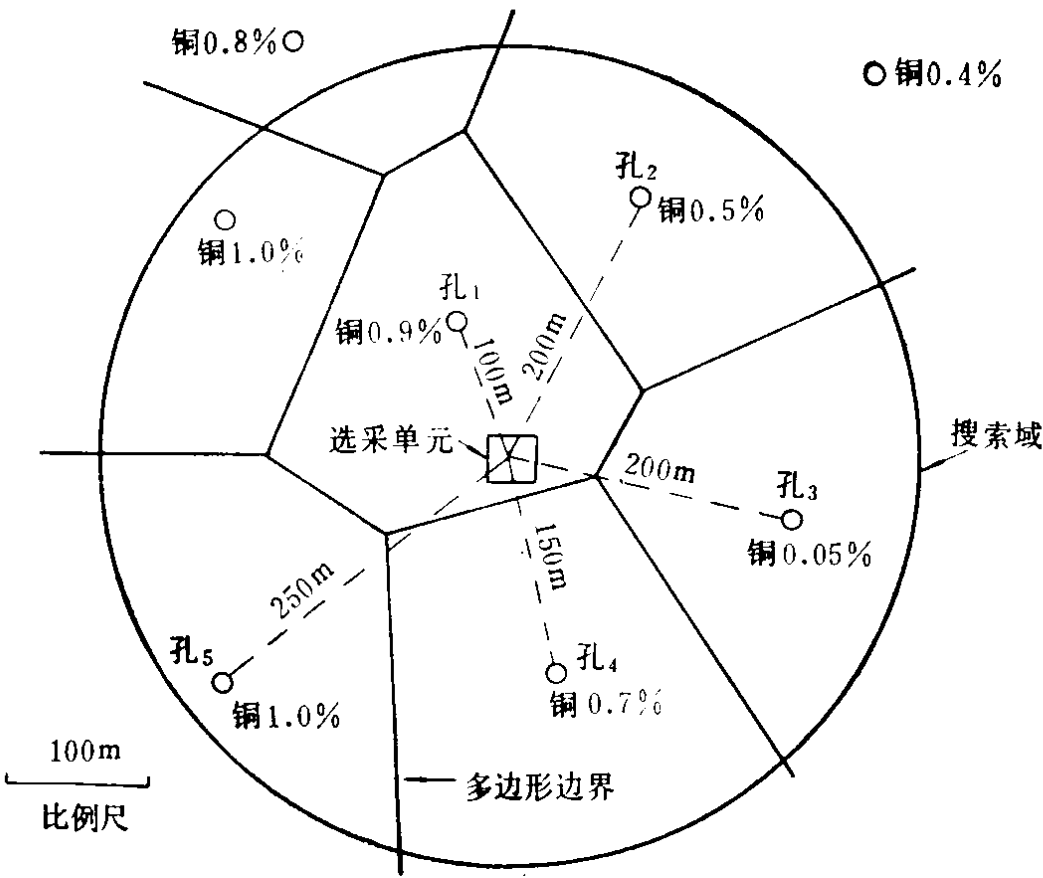
上述问题说明：传统储量计算方法在估值精度上可能存在系统的偏差，因此往往不能适应矿床开发中精确计算储量和误差评价的要求。

地质统计学方法（又称克里格法）用于储量计算时，充分利用了已知信息，考虑到不同方向上变异性的差异，在统一支集的基础上求算真值的最佳无偏估计量，因而其估值具有较高的精度，这正是地质统计估值的优点（图1-1-3）。

4) 经典统计学方法对矿床地质变量的研究

用确定性解析式为数学模型，在储量计算方法上不断演进的同时，用经典统计学方法对矿床地质变量的研究也在进行。在长期的地质、矿业实践中，矿床地质变量具有随机性特点已逐渐为人们所认识。加之，矿床勘探的基本方式是对矿床的离散化取样，再由离散取样的观测值去推断矿床的赋存特征，这套工作方法从形式上看与数理统计学的研究方法十分相似。因此，很早就有人从经典统计学角度对矿床地质变量进行了研究。

早在30年代初期，苏联的地质勘探人员就已经开始用数理统计方法研究矿床的变化性、勘探网度和储量误差三者之间的关系。他们用简单的区间估计公式计算勘探网度和储量



钻孔号	权系数			
	多边形法	距离反比法	距离平方反比法	克里格法
1	1.000	0.326	0.475	0.471
2	0.000	0.163	0.119	0.058
3	0.000	0.163	0.119	0.128
4	0.000	0.218	0.211	0.256
5	0.000	0.130	0.076	0.088
块段估计 铜品位	0.90%	0.74%	0.77%	0.78%
预计精度	± 0.697	± 0.568	± 0.560	± 0.557
估计方差	0.486	0.322	0.314	0.310

图 1-1-3 各种估值方法的比较

估计误差。在50~60年代，有的苏联学者专注于对煤田有关地质变量（如灰分、硫分）统计分布的研究，并得出灰分基本上是正态分布等结论。但是很快发现，这样做很难符合地

质实际，因为很多地质变量虽然具有随机变量的特点，但又不是纯粹的随机变量。

本世纪40年代，南非学者H.S.西舍尔根据对南非兰德金矿手工采样的研究，首先提出了金品位的对数正态分布模型，给出储量计算用的公式和图表，并估计采样时可能出现的误差。随后D.G.克里格又提出了三参数对数正态分布模型。从此，以统计学方法评价金矿床的工作得到迅速的推广。他们在南非的工作曾因此而取得相当令人瞩目的成绩，但是把经典统计学方法简单地用于地质变量研究，特别是企图以此来解决复杂的地质问题是不可能的，所以这样做不会总是成功的。

用经典统计学方法研究矿床地质变量，至少存在以下几方面的问题：

① 经典统计学研究的对象是随机变量，而地质变量的取值虽然具有随机变量的特点，但又不是纯粹的随机变量。地质与矿业上遇到的很多地质变量是一种既有随机性又有结构性（指在空间分布上所具有的相关性）的变量。

② 经典统计学的采样是独立的，即要求采样样本中的各个采样点及其观测值（样品值）是互相独立的，不考虑样品的空间分布。地质勘探中对矿床的采样正相反，采样点的布置必须充分考虑它们的空间分布，相邻样品由于空间的接近在取值上具有相关性，因此采样不是独立的。

③ 经典统计学的研究是建立在随机现象可以无限次重复试验或观测的基础之上的，要求每次试验或观测具有相同的条件。这种要求用在对地质变量的观测上则不行。对于一个矿床（如煤层）来说，在其某一位置上进行钻探取样后，就不再可能在同一位置上进行再次取样了。

④ 经典统计学方法不可能反映矿床在空间上的变异规律，只能得出不考虑空间变化规律的纯统计规律。这一缺陷对经典统计学用于地质变量研究是致命的。也就是说，经典统计学研究地质变量具有很大的局限性。如图1-1-4(a)、(b)、(c)、(d)中，4个钻孔所见煤层的厚度都是1、2、3、4 m。虽然它们的平均厚度和方差是一样的，但由于钻孔位置不同，它们的厚度变化情况很不相同。

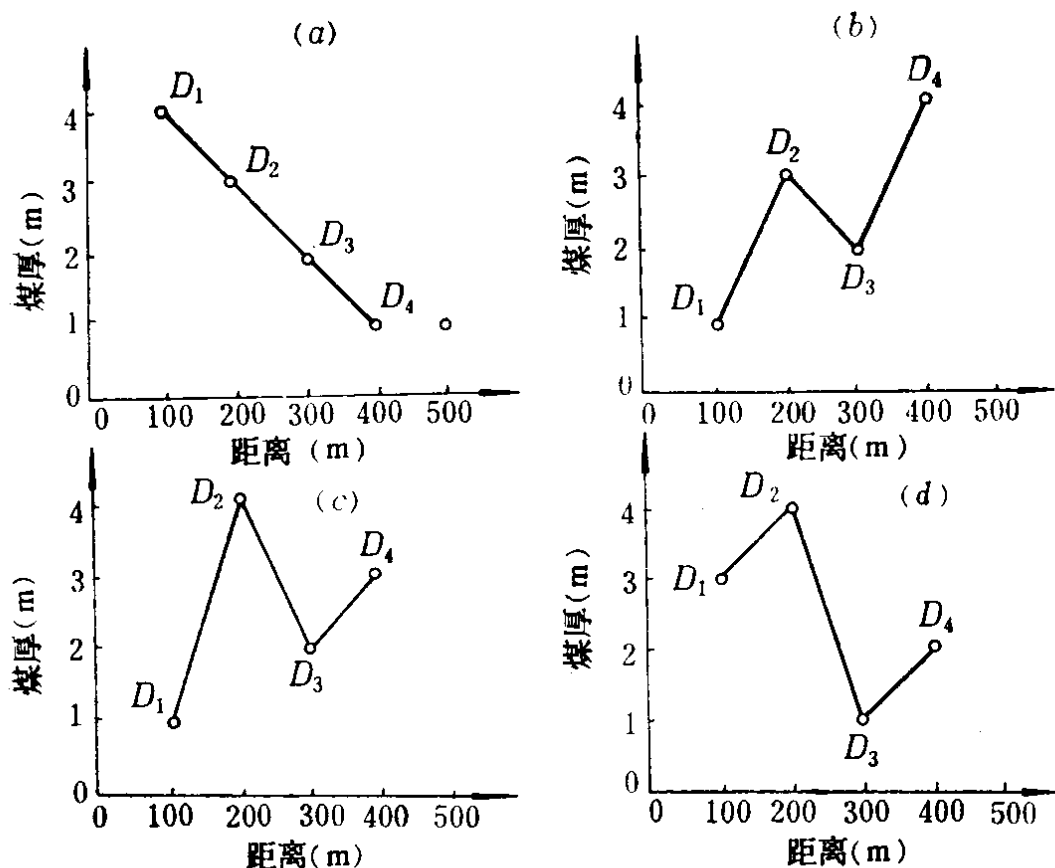


图 1-1-4 煤层厚度的均值及方差相同但空间变化规律却明显不同的典型示例

$D_1 \sim D_4$ 表示不同位置和煤厚的钻孔

从70年代开始，我国煤田勘探系统普遍用煤层厚度的变异系数来表示煤层厚度的稳定性大小。由于这个变异系数是