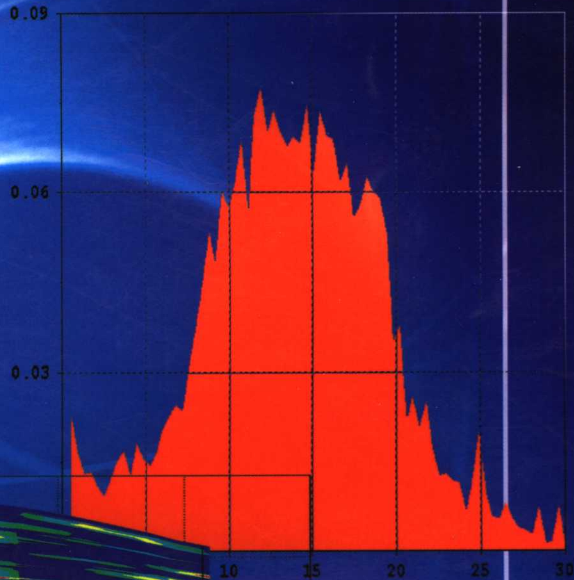
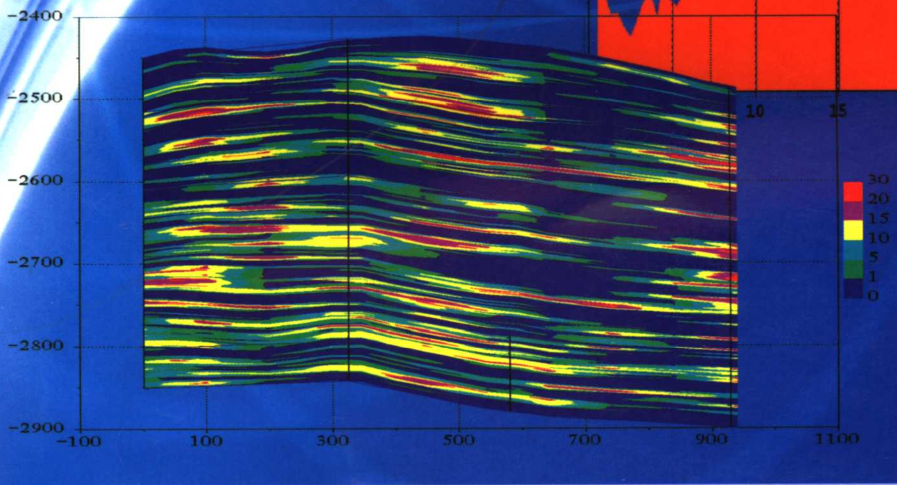
高等学校教材

现代数学地质

Modern Mathematical Geology

康永尚 沈金松 谌卓恒 编著





石油工业出版社

高等学校教材

现代数学地质

康永尚　沈金松　谌卓恒　编著

石油工业出版社

内容提要

本教材是针对油气勘探开发人才的培养而设计的，内容覆盖面广，从概率论和数理统计的基础知识到其应用，从单变量分析到多变量分析，从空间无关的变量到空间相关的变量，从线性科学的思维观到非线性科学的思维观，从易到难渐次深入，构成一个油气地质和开发领域从基础到应用的数学地质方法体系。

本书可作为高等院校地质勘探、地质工程及环境工程相关专业本科生和研究生的专业基础课教材，也可供石油地质等专业的生产和科研单位相关人员参考。

图书在版编目(CIP)数据

现代数学地质/康永尚等编著.
北京:石油工业出版社,2005.6
高等学校教材
ISBN 7-5021-5097-8

Ⅰ.现…
Ⅱ.康…
Ⅲ.数学地质-高等学校-教材
Ⅳ.P628

中国版本图书馆CIP数据核字(2005)第051970号

现代数学地质
康永尚　沈金松　谌卓恒　编著

出版发行:石油工业出版社
(北京安定门外安华里2区1号　100011)
网　址:www.petropub.com.cn
总　机:(010)64262233　发行部:(010)64210392
经　销:全国新华书店
印　刷:北京华正印刷厂

2005年6月第1版　2005年6月第1次印刷
开本:787×1092毫米　开本:1/16　印张:11.75
字数:290千字　印数:1—2000册

定价:16.80元
(如出现印装质量问题,我社发行部负责调换)

前　言

地质学最重要的应用是各种自然资源的开发和利用，随着社会的发展，越来越多的数据资料不断积累，从这些数据中挖掘规律和知识，对于更好地开发和利用各种自然资源具有重要的意义。欧洲经济共同体一位负责人曾经说过一句非常深刻的话："欧洲的许多矿产资源存在于数据库中"，我们可以这样理解：通过矿产数据的分析，揭示在数据中存在的信息，可以对矿产赋存特点有进一步的认识，进而帮助我们找到更多的资源。

无论是作为勘探地质工作者，还是作为油气藏工程师，都越来越强调对地质现象或地质体定量研究的能力，尤其是随着资料的不断积累，这种趋势在不断加强。

利用数据资料形成认识的过程是一个分析和归纳的过程，而利用这种认识指导实践的过程是一个推断过程，前者是途径，后者是目的，认识世界最终是为了改造世界。

定量分析的基础是数据，分析过程是建模过程，建模有确定型和随机型两种模式。确定型模式探求的是某一变量与其影响因素之间的函数关系，这当然是科学研究应当追求的最终模式，因为，函数关系具有逻辑和形式上的美学意义，它带来的是精确的预测，是人们对研究变量百分之百的掌控。然而，自然现象不同于人类可以操控的室内物理实验，任何一个自然变量都是在"上帝"操控下的实验结果，尤其是地质变量，不仅有"无数只手"在影响它、改变它，而且，这种影响和改变历经漫长岁月，使得变量与其影响因素之间的关系变得极其复杂，以至于建立函数关系的企图遭遇严峻的挑战。

然而，复杂中孕育着简单，大量的数据表现出规律，也正是如此，统计学这一基本的工具在地质现象的研究中得到了广泛的应用。显然，用统计学这一工具研究地质现象，体现的正是人类的智慧和不屈不挠的科学精神，数学地质也正是在这一基础上不断发展成为一个庞大的知识系统。

现代数学地质涉及地质学的诸多领域，对于油气地质工作者来说，究竟需要掌握哪些内容，也就是说，一本适合的石油高等院校教材应包括哪些主要内容？这是我们长期思考的一个问题。在中国石油大学的教学工作为我们提供了一条探索这一问题的有效途径。

我们从 1994 年开始，先后在中国石油大学（北京）和中国石油大学（华东）主讲了研究生（包括进修生）和博士生课程《地质统计学》、《分形几何》、《数学地质》和《现代数学地质》，累计 20 次之多，对研究生层次的数学地质培养进行了有益的探索，在空间相关统计学和分形几何方面，进行了理论方法和应用实例方面的总结，同时，相关方面的科研工作也起到了一定的推动作用，使这部分内容逐步得到了充实和完善。

本科生数学地质的教学实践使我们回过头来，考虑更加基础的内容——概率论、数理统计、线性回归和多元统计分析。教材在体现这些内容时试图从简单的实例出发引出定理和公式。

希望这本教材能引导更多的读者对数学地质产生兴趣并在今后的工作中自觉地应用数学地质方法，促进研究工作向半定量化和定量化方向发展，为油气勘探开发事业做出更大的贡献。

教材的编写和顺利出版是多方面持续支持和努力的结果，我们首先要感谢中国石油大学

(北京)资源与信息学院书记庞雄奇教授、院长朱筱敏教授和副院长陈小宏教授和钟宁宁教授对该教材出版的大力支持。我们感谢资源与环境系和盆地与油藏研究中心的领导与同事们对我们教学工作的支持和帮助。同时,我们要特别感谢彭仕宓教授对教材编写工作的支持和尹志军博士对教材编写的贡献。

中国石油大学张一伟教授和熊琦华教授长期致力于统计学方法在油藏定量描述中的应用,有力地推动了现代数学地质在石油大学的教学和科研应用工作,本教材也凝聚着前辈的心血,在此,对他们的杰出工作表示崇高的敬意。这里还要提及的是金之钧教授首先提出在中国石油大学(北京)为研究生开设数学地质课程的计划并编写了教学大纲,虽然因其有其它更重要的工作而未讲授该课程,但他的努力为我们后来开设数学地质课程奠定了基础,也为本教材的形成提供了教学实践契机,在此表示感谢。

本教材前言、绪论、第一章、第二章的第一节到第四节中的参数估计部分以及第五章到第七章由康永尚编写,其中,第七章中随机模拟应用实例由尹志军博士提供;第八章以康永尚编写为主,沈金松补充了部分应用实例;第二章第四节中的假设检验部分以及第三章到第四章由沈金松编写,在第四章编写过程中吸收了旅居加拿大的谌卓恒博士提供的素材和应用实例;全书由康永尚统稿。研究生杨帆、徐显生和徐桂华等协助完成了教材中部分图件的清绘和数据表的制作及其它事务性工作,对他们的工作表示感谢。

中国石油大学(北京)陈安乐教授对全书进行了详细的审校,提出了大量宝贵意见,在此,编著者衷心感谢陈老师的辛勤劳动。

本教材是编著者长期教学实践的总结和凝练,编著者的家庭给予的支持是完成这一工作最有力的保证,在此表示感谢。

尽管几经修改,但教材中的错误在所难免,欢迎并感谢读者能给予指正并与我们取得联系,以便该教材体系能在未来的实践中不断完善。我们的联系方式是:kangysh@sina.com,jinsongshen@hotmail.com,zchen@shaw.ca。

编著者

2005 年 4 月

目　　录

绪　　论

《现代数学地质》学习目的和要求：能够把概率论和数理统计知识熟练地应用到油气勘探开发的决策分析和地质变量变化规律的研究和预测中，掌握多元统计学的基本原理及其应用，深刻理解空间相关统计学的精髓并能应用于空间变量的定量化研究中，了解非线性科学的新思维和分形几何的基本原理及其应用。

一、数学地质的起源与发展

1.数学地质的定义

关于数学地质的定义，尚未有一致的说法，Davis(1973)把数学地质定义为“地质数据的定量分析方法”，而 Agterberg(1974)把数学地质定义为“地球科学中的全部数学应用”，赵鹏大(1983)认为“数学地质是研究最优数学模型并查明地质运动数量规律性的科学”。无论采用何种定义，我们都可以把数学地质理解为以数学作为基本的理论和方法，以计算机作为工具，对地质现象进行定量化研究的一门边缘学科，是由地质学、数学和计算机科学互相结合而逐渐发展完善的。

2.数学地质产生的原因及对地球科学的影响

传统的地质学基本上是一门描述性的学科，从事地质学研究的人对数学在地质学研究中的应用及所起的作用还存在不同的看法。然而，随着人类对地球研究和探索的深入，需要研究多种地质因素对某一地质现象的影响，分析多种地质因素间的相互关系及各自的相对重要性，对于这样的一系列复杂问题没有数学作为工具是不行的。其次，随着研究手段的多样化，通过实验和观察获得了大量的数据，要利用这些数据研究各种量之间的关系及各种量的变化等反映地质运动和变化的规律，必须应用数学方法和计算工具。再次，数学是研究和模拟各种地质现象的理论基础，大气运动规律、地壳变化和运动的研究等都需要数学模拟。

随着数学向传统地质学的渗透和应用，地质学研究逐渐由定性向定量分析过渡，也由最初的单变量分析向多变量分析过渡。所研究的对象也由确定性事件向随机性事件过渡，对地质事件和地质现象的分析也由定性解释向定量分析与数值模拟相结合的阶段发展。

3.数学地质发展的 4 个阶段

数学被逐步引入到地质学中经历了 160 多年的历史，作为一门新兴的边缘学科，数学地质已有 60 多年的发展历史。从应用的广泛性看，数学地质的发展大致经历了以下 4 个阶段。

1)孕育阶段(1950 年以前)

1840 年英国地质学家 Lyell 首次以古生物化石的统计分析为论据，对第三纪地层进行了划分，确定了岩石地层次序，著述了《定量动物学》一书，开创了数学方法引入地质学问题研究的先例。上世纪 30 年代后期，Simpson(1939) 著述的《分析地质学》一书中列举了统计学在生物研究中的多方面应用。Burma(1949)在“多元分析——地质学和古生物学中的一种模型分析工具”一文中明确提出了多元统计方法是一种最有前景的生物计量方法。Krumbein 和格里菲斯在沉积学研究中也使用了概率统计方法，为用数学方法研究地质现象和个别指标的统计分析解决具体地质问题奠定了基础。这一阶段也是统计分析在地质学中应用的可能性问题讨论最

为激烈的时期。

2)早期阶段(1951～1960)

Krumbein(1956)在研究岩石的矿物成分、岩性和化学成分时,应用了多元统计方法,并把岩石成分作为 n 维空间中的一个点或向量进行统计处理。尤其值得指出的是,1958 年由 Krumbein 和斯洛斯公开发表了第一个面向地质应用的计算机程序,标志着计算机技术在地质研究中应用的开端,加速了地质学的定量化研究进程。1958 年 Sichel 和 Kridge 编著的《地质统计学》及 Allais(1957)发表的《单元中矿床数服从泊松分布的矿产资源定量评价》等重要文献奠定了地质统计学分支学科的基础,是数学地质早期发展的重要阶段。

3)形成发展阶段(1961～1980)

自 1968 年在布拉格第 23 届国际地质大会上成立国际数学地质协会(IAMG)并由维斯捷列马斯任第一届国际数学地质协会主席以来,数学地质这一边缘交叉学科得到了长足的发展。1969 年出版了两种专门的数学地质刊物——《Mathematical Geology》和《Computers & Geosciences》,后来又设立了新刊物《Nonnewreable Resources》。随着人类对资源需求的增加,社会对资源评价和预测研究精度的要求也逐渐提高,迫使地质学家和工程技术人员寻求更加可靠和精确的评价和预测模型。例如,在石油禁运和社会对油气需求增大的背景下,Harris(1973)发表了多元统计评价及主观概率评价两种油气资源评价模型。为了加强矿产资源的定量评价,Agterberg(1974)针对法国和非洲的一些固体矿床的储量和品位的评价,提出了矿产资源评价的逻辑模型。

4)深入广泛的发展阶段(1980 年至今)

这一阶段中,数学地质向更广泛和更高水平发展。随着超大规模集成电路的计算机的研制成功和各种新的应用数学模型的建立,促进了数学地质向更加深入和稳健的方向发展。如日本林知已夫的数量化理论,美国麦克卡门和波特波尔的特征分析和前苏联的康士坦丁诺夫的逻辑信息方法等为定性地质变量和定量地质变量的联合数学模型的建立提供了基础。地质过程的计算机模拟、地质数据库的建立和地学领域内人工智能专家系统的研制和应用,为数学地质的广泛应用提供了更加先进和方便的手段。

由于计算机和计算数学的发展,新的多元统计分析方法和新的数学理论和方法不断与地质结合,使数学地质更加完善和成熟。人们开始研究一些新的数学地质方法,如法国著名的数学家 Tom(1968)提出了用数学工具描述灾难性或突如其来的变化现象的突变理论,后来又将微分拓扑学的研究成果应用于地质学,建立了地质学中的突变理论模型,用于研究断层运动、二叠纪海洋无脊椎动物的灭绝原因等。

人工智能及专家系统的研究使专家的知识和经验能为他人所用,充分发挥了计算机和专家的作用;自美国斯坦福国际人工智能研究所于 1976 年研制成功地学领域内的第一个专家系统——Prospector 以来,利用该系统美国的一家勘探公司已在华盛顿州发现了一个钼矿床。美国的石油资源评价专家系统及斯仑贝谢跨国测井公司研制的地层倾角测井资料处理解释咨询系统(Dipmeter Advisor)已在油气勘探和开发中取得了很好的效果。地质数据库的成功应用为资料的交流和综合应用提供了可能。

二、数学地质的最新进展

从近年数学地质发展的趋势看,数学地质正全方位地向各地质分支学科渗透,数学地质与新发展的其它学科交叉的深度和广度在不断扩大,较为突出的表现为以下几个方面:

(1)非线性科学在地质学中的应用。

自变量的连续改变而发生不连续变化的数学理论——突变理论的发展。突变现象的共同特点是，条件微变会导致系统宏观状态的剧变。例如，地质学中的突变理论模型；用于研究断层运动、二叠纪海洋无脊椎动物的灭绝原因等方面取得了可喜的进展。

(2)人工智能在地质学中的应用：专家系统在地质学中的应用；多种人工智能技术在地质学中的综合应用，如神经网络、遗传算法等。

(3)地理信息系统：在计算机软件的支持下，空间数据输入、存储、检索、运算、显示和综合分析的应用技术系统在地学领域中的应用。

(4)数据库技术在地质学中的应用。

(5)自动绘图技术在地质学中的应用。

(6)计算机的最新硬软件在地质学中的应用。

从数学地质的最新进展来看，“非线性科学”认识论在理论分析方面的开拓和已成熟方法的软件化，代表两个主流。

三、现代数学地质包含的主要内容

根据以上对数学地质定义的理解，我们可以说数学地质是一门迅速发展、开放的学科。随着计算机技术的不断提高和地球科学的发展，计算机应用的日益普及，地质现象的定量化研究已成为一个必然趋势，数学地质要解决的问题正是地质现象的定量化这一极为困难的问题。作为一门把地质问题定量化不断推向深度和广度的、具有生命力的综合性边缘学科，正蓬勃发展着，它以多种数学方法为基础，同时吸收了现代物理学、化学、社会科学等研究成果，借助于计算机，实现对地质过程的模拟、对地质现象时空变化的分析、对多种变量之间相互关系的刻画，而以各种预测为最终目标，为人类的生产实践提供决策的量化依据。《现代数学地质》包括以下主要内容：

(1)数据分布类型及预处理：地质数据的分布类型及其检验数据的特征分析，变量的选择及最佳组合，地质数据之预处理；

(2)属性分析与综合：相关、聚类、判别、对应、因子分析等；

(3)地质多元统计分析方法：趋势面分析、特征的空间分布、单位向量场分析；

(4)矿产资源统计预测：Delphi 法、主观概率、蒙特卡罗、条件概率等；

(5)时间序列分析预测；

(6)地质统计学；

(7)模糊数学和突变论；

(8)分形理论；

(9)灰色系统；

(10)地质过程的计算机模拟(如盆地模拟)；

(11)地质数据库及地质数据处理系统(如 GIS)；

(12)人工智能专家系统。

现代数学地质是一个庞大的知识体系，根据我们的教学经验，针对油气勘探开发地质和工程人才培养的需要，试图建立起从数理统计基础到专业应用的一个实用且相对完整的知识结构体系并努力使本教材具有如下特色：

(1)针对性：主要针对石油勘探开发的需求设计教材的内容。

(2)新颖性：涉及风险分析、空间相关统计学、非线性科学思维观和分形几何等方法，有些方法在国内目前应用还比较少，但随着我国社会经济的发展、资料数据的积累尤其是油公司体制的逐步确立，其应用前景良好。

(3)系统性：从概率论基础开始，逐步深入，过渡到空间相关统计学比较前沿的领域，最后，引入非线性科学的思维观和方法论。

(4)学科交叉性：软硬学科交叉(决策科学和地质变量定量表征)、勘探和开发的学科交叉。

(5)实用性：每部分方法原理之后，附有在勘探开发领域内的应用实例。

(6)应用的广泛性：适用于“地质工程”、“矿产普查与勘探”、“地球探测与信息技术”、“油气井工程”、“油气田开发工程”和“环境科学”等多个二级学科。

根据以上思路本教材设计了 8 章内容，从概率论和数理统计的基础知识到其应用，从单变量到多变量，从空间无关的变量到空间相关的变量，从线性科学的思维观到非线性科学的思维观，从易到难渐次深入：第一、第二章作为概率论和数理统计的基础知识，从基本概念、基本公式、基本方法和在油气勘探开发中的应用几个方面进行了介绍和讨论；第三章为线性相关和线性回归，其重点是阐明最小二乘法的基本原理及简单和多元线性回归方程的建立与应用；第四章为多元统计学，主要介绍了趋势面分析、主成分分析、聚类分析和判别分析等多变量研究中常用的方法；第五章到第七章为空间相关地质变量研究——地质统计学的主要内容，从区域化变量的结构分析、普通克立金、指示克立金和地质变量的随机模拟及其应用等方面进行了分析讨论；第八章涉及非线性科学的基本思维观和分形几何的基本原理及其应用。

第一章　概率论基础及其在油气勘探开发决策中的应用

本章学习目的和要求：简要回顾概率论基本知识，介绍概率论在油气勘探风险决策分析中的应用，为以后各章的学习奠定基础，并使学生能熟练地把概率论的基本方法应用到油气勘探开发的决策分析中。

第一节　概率论基础

一、基本概念

在自然界和人类社会中存在着两类不同的现象，一类是确定性现象，另一类是随机性现象。

1.确定性现象和随机性现象

在一定条件 S 下进行某种试验或观察，事先可以预料必然出现某种结果，或者说在条件 S 不变的情况下，重复进行一系列这种试验或观察，所得结果是完全相同的。像这类在一定条件下必然出现某种结果的现象称为确法性现象，如纯水在一个大气压下加热到100℃时必然沸腾，异性电荷必然互相吸引、同性电荷必然互相排斥等皆为确定性现象。

在一定条件 S 下进行某种试验或观察，在试验或观察之前无法预知确切的结果，只知道可能出现这样的结果，也可能出现那样的结果。当然，一旦试验或观察进行之后，出现什么结果就是完全确定的了。但是，在试验条件 S 不变的情况下，重复进行一系列这种试验或观察所得的结果却是不尽相同的。像这类在一定条件下可能出现的结果不止一个，至于出现哪一个，事先又无法确定的现象称为随机性现象。如某一地区在某一天的降雨量，电话交换台在单位时间内收到的呼唤次数等皆为随机性现象。

2.随机试验和样本空间

自然界和人类社会中各种各样的现象常常在人们所进行的试验或观察中呈现出来，我们把呈现出随机性现象的试验或观察称为随机试验，记为 E。

在进行一个随机试验时，试验的所有可能结果应该是明确知道的，每一个结果被称为样本点，用 ω 表示。由样本点 ω 的全体构成的集合称为样本空间，用 Ω 来表示，记为 $\Omega=\{\omega\}$。

【例1-1】　掷一枚骰子，试验的所有可能结果有6个，即 $\omega_1=$“1”，$\omega_2=$“2”，$\omega_3=$“3”，$\omega_4=$“4”，$\omega_5=$“5”，$\omega_6=$“6”，则样本空间为

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$$

【例1-2】　观察电话交换台在单位时间内收到的呼唤次数，若用 k 表示“在单位时间内收到 k 次呼唤”这一结果，$k=0,1,2,\cdots$，则样本空间为

$$\Omega=\{0,1,2,\cdots\}$$

3.随机事件

在一个随机试验中,对一次试验而言,可能出现也可能不出现的事情,称为随机事件,记为 A。对一个随机试验来说,显然,它的每一个可能的结果(即每个样本点)都是一个随机事件,我们称之为基本事件,像例 1-1 中出现 6 点的事件即为一个基本事件。在一个随机试验中,除了基本事件外,还有由若干个可能结果(即若干个样本点)所组成的事件,相对于基本事件,称这种事件为复合事件,如例 1-1 中出现偶数点的事件即为一个复合事件。

4.频率与概率

在相同条件 S 下,重复进行 n 次试验,若在 n 次试验中,事件 A 发生的次数(可称为频数)为 μ_A,则比值$\frac{\mu_A}{n}$称为事件 A 在 n 次试验中发生的频率 f,即

$$f = \frac{\mu_A}{n} \tag{1-1}$$

如果当 n 充分大时,A 发生的频率$\frac{\mu_A}{n}$稳定地在某一数值 p 附近摆动,而且一般来说随着试验次数的增多,这种摆动的幅度越变越小,即频率$\frac{\mu_A}{n}$越来越稳定于 p,则称 p 为此随机试验中随机事件 A 发生的概率,记作

$$P(A) = p \tag{1-2}$$

【例 1-3】 表 1-1 为 17 块岩心的实测岩石密度数据,假设事件: $A_1 = [2.65,2.68]$;$A_2 = [2.68,2.71]$;$A_3 = [2.65,2.71]$;$A_4 = [2.66,2.69]$,计算各事件的概率。

解: $P(A_1) = \frac{7}{17} = 0.412$;$P(A_2) = \frac{9}{17} = 0.529$;$P(A_3) = \frac{16}{17} = 0.941$;

$P(A_4) = \frac{12}{17} = 0.706$

表 1-1 实测岩石密度数据表

岩心号	岩石密度 ρ,g/cm^3	岩心号	岩石密度 ρ,g/cm^3	岩心号	岩石密度 ρ,g/cm^3
1	2.68	7	2.69	13	2.69
2	2.68	8	2.70	14	2.68
3	2.68	9	2.68	15	2.68
4	2.69	10	2.69	16	2.68
5	2.69	11	2.70	17	2.70
6	2.70	12	2.74		

5.事件之间的关系与运算

1)包含关系

若事件 A 发生时,事件 B 必发生,则称事件 B 包含事件 A,见图 1-1(a),记为 $A \subseteq B$。例 1-3中的 A_3 包含事件 A_1,同时,A_3 也包含事件 A_2。

2)事件的和

若某事件发生当且仅当事件 A 与事件 B 中至少有一个发生(或者说当且仅当事件 A 发生或事件 B 发生),则称此事件为事件 A 与事件 B 的和见图 1-1(b),记为 $A + B$ 或 $A \cup B$ 。

3)事件的积

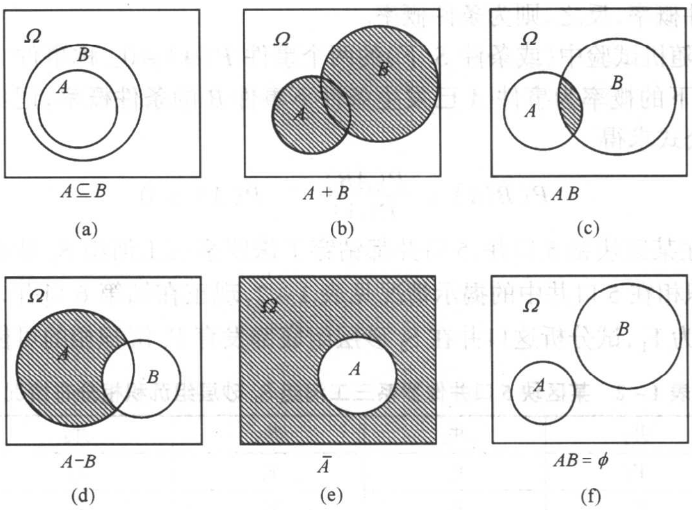


图 1－1　事件之间关系及运算示意图

若某事件发生当且仅当事件 A 与事件 B 同时发生，则称此事件为事件 A 与事件 B 的积，见图 1－1(c)，记为 AB 或 $A \cap B$。例 1－3 中事件 A_1 和 A_4 的积(同时发生)为事件[2.66,2.68]。

4)事件的差

若某事件发生当且仅当事件 A 发生而事件 B 不发生，则称此事件为事件 A 与事件 B 的差，见图 1－1(d)，记为 $A - B$。

5)对立事件

事件“A 不发生”为事件 A 的对立事件，记为 $\bar{A}$，见图 1－1(e)。

6)互不相容事件

若事件 A 与事件 B 不能同时发生，则称事件 A 与事件 B 为互不相容事件，见图 1－1(f)。例 1－3 中 A_1 和 A_2 两个事件即为互不相容事件。

6.概率的可加性

1)概率的有限可加性

若事件 $A_1, A_2, \cdots, A_n$ 两两互不相容，则

$$P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{1-3a}$$

式(1－3a)也可写成如下形式

$$P(\sum_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) \tag{1-3b}$$

2)概率的可列可加性

若事件 $A_1, A_2, \cdots, A_n \cdots$ 两两互不相容，则

$$P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \tag{1-4}$$

二、条件概率公式和概率的乘法公式

1.条件概率公式

随机事件 A 的概率是在条件 S 之下来考虑的，若除条件 S 之外，不再附加任何条件，则称

这种概率为无条件概率，反之，则为条件概率。

设 A、B 是某随机试验中(或条件 S 下)的两个事件 $P(A)\neq 0$。称事件 B 在“事件 A 已发生”这一附加条件下的概率为事件 A 已发生条件下事件 B 的条件概率，记 $P(B/A)$。条件概率可由条件概率公式求得

$$P(B/A)=\frac{P(AB)}{P(A)} \qquad P(A)>0 \tag{1-5}$$

【例1-4】 在某区块钻5口井，5口井都钻穿了侏罗系三工河组 S_2 砂层组，该砂层顶部、中部和底部的沉积相在5口井中的揭示情况见表1-2，现正在钻第6口井，已钻到 S_2 砂层组的中部，其沉积相为 F_1，试分析这口井在 S_2 砂层组底部发育 F_3 沉积相的可能性有多大？

表1-2 某区块5口井侏罗系三工河组 S_2 砂层组沉积相分布情况

位置	井$_1$	井$_2$	井$_3$	井$_4$	井$_5$
顶部	F_3	F_2	F_1	F_2	F_1
中部	F_1	F_1	F_2	F_1	F_2
底部	F_2	F_3	F_3	F_3	F_3

解：设事件 $A=[F_1$ 是 S_2 砂层组中部的沉积相$]$，事件 $B=[F_3$ 是 S_2 砂层组底部的沉积相$]$，那么，要回答的问题是当事件 A 发生时，事件 B 发生的条件概率。

根据表1-2的信息，我们有

$$P(A)=\frac{3}{5},P(B)=\frac{4}{5}$$

事件共同发生的概率为 $P(AB)=\frac{2}{5}$，则在事件 A 已发生的条件下事件 B 发生的概率为

$$P(B/A)=\frac{P(AB)}{P(A)}=\frac{2/5}{3/5}=\frac{2}{3}$$

故有：当第6口井 S_2 砂层组的中部沉积相为 F_1 时，在 S_2 砂层组底部发育 F_3 沉积相的可能性为2/3。

2.概率的乘法公式

由式(1-5)可得概率的乘法公式

$$P(AB)=P(A)P(B/A) \tag{1-6}$$

在关系式 $P(AB)=P(A)P(B)$ 成立的前提下，称事件 A 和 B 相互独立。

注意，在讨论两个事件 A 和 B 的独立性时，A 和 B 必须是相容事件，否则，就无从谈其独立与否。

在例1-4中，$P(AB)=\frac{2}{5}\neq P(A)P(B)=\frac{12}{25}$，故 A 和 B 两个事件不是相互独立的，也就是说，在 A 发生时，对 B 发生的概率会产生影响。

三、全概率公式和概率树

1.全概率公式

在实际问题中，我们常常遇到这样的事件 B，它总是与若干个两两互不相容的事件 A_1，$A_2\cdots A_n$ 之一同时发生。这时，为了计算事件 B 的概率，可先把事件 B 分解为一些不相容的简单事件之和，即 $B=BA_1+BA_2+\cdots+BA_n$，然后分别计算和式中的简单事件的概率，再利用概率的可加性，最终得到 $P(B)$。下面，先从实例分析入手，引出全概率公式。

【例 1-5】 某油层中有三种沉积微相,已知其中第一、第二和第三种沉积微相各占总量的 50%、25%和 25%,又知第一、第二和第三种沉积微相高渗透($>100\times10^{-3}\mu m^2$)岩心出现的概率分别为 1%、2%和 4%,现任取一个岩心,问取出的岩心具有高渗透率的概率有多大?

解:设 A_i =“取出的岩心是第 i 种沉积微相”,则有

$$P(A_1)=50\%,\quad P(A_2)=25\%, P(A_3)=25\%$$

再设 B =“取出的岩心是高渗透率”,事件 B 总是与三个不相容的事件 A_1、A_2 和 A_3 之一同时发生,且 $P(B/A_1)=1\%$, $P(B/A_2)=2\%$, $P(B/A_3)=4\%$,根据全概率公式,事件 B 的概率为

$$P(B)=\sum_{i=1}^{3}P(A_iB)=\sum_{i=1}^{3}P(A_i)P(B/A_i)$$
$$=50\%\times1\%\times25\%\times2\%\times25\%\times4\%=2\%$$

显然,事件 C =“取出的岩心不是高渗透率”的概率为

$$P(C)=1-P(B)=1-2\%=98\%$$

由以上实例分析,可引出全概率公式如下:

设:(1) $A_1, A_2, \cdots, A_n$ 两两互不相容, $P(A_i)>0\quad(i=1,2,\cdots,n)$

(2) $B\subseteq\sum_{i=1}^{n}A_i$(事件 B 包含在 $\sum_{i=1}^{n}A_i$ 中)

则有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B/A_i)=\sum_{i=1}^{n}P(A_iB)\qquad(1-7)$$

2.概率树

我们可以用树形结构图清晰地表示事件之间的概率关系,例 1-5 中事件之间的概率关系可用图 1-2 所示的概率树来表示。

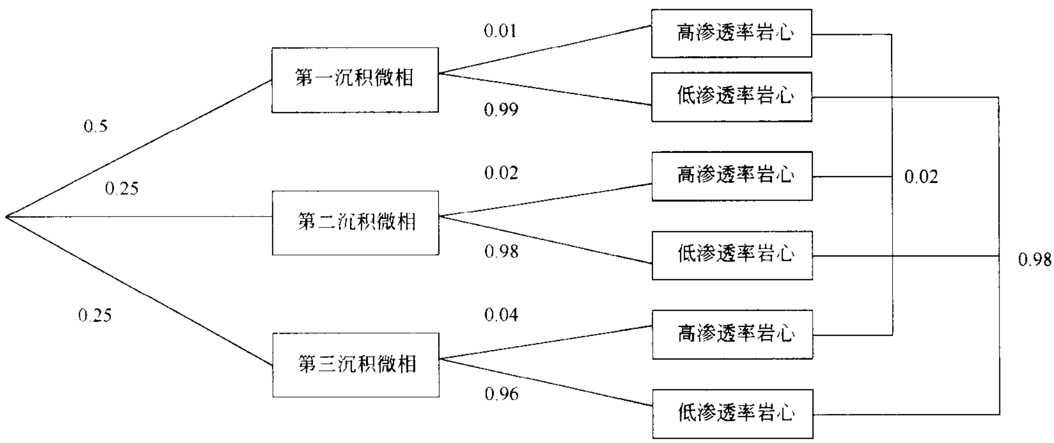


图 1-2 例 1-5 中事件之间概率关系的概率树

四、贝叶斯公式和逆概率树

1.贝叶斯公式

在实际工作中还常常碰到这样一类问题:已知某个试验结果是由许多原因导致的,如果人们通过试验观察到这个结果,于是人们希望通过这个信息来探讨(反求)每个“原因”导致这个结果的可能性有多大,如例 1-4 中,若已知取出的一个样品为高渗透样品,问这个样品属于第

一、第二和第三种沉积微相的可能性有多大？由此引出贝叶斯问题，贝叶斯问题实际上是逆概率问题，在决策过程中有十分重要的应用价值，贝叶斯公式由下列方式表述：

若有一组事件 $A_1, A_2 \cdots A_n$ 以及事件 B，它们满足公式(1－7)中相同的条件，而且 $P(B) > 0$，则有贝叶斯公式

$$P(A_i/B) = \frac{P(A_i)P(B/A_i)}{\sum_{j=1}^{n} P(A_j)P(B/A_j)} = \frac{P(A_iB)}{P(B)} \quad (j = 1,2,\cdots,n;\quad i = 1,2,\cdots,n) \quad (1-8)$$

实际上，式(1－8)的贝叶斯公式可根据条件概率公式和全概率公式直接推导出来。

由式(1－8)所计算的概率称为后验概率，它实质上是 B 发生条件下 A_i 发生的条件概率，这不同于 A_i 本身发生的概率，称为先验概率。

下面，用一个例子来说明贝叶斯公式的应用。

【例 1－6】 回到刚刚提出的问题上，即在例 1－5 中，若已知取出的一个岩心样品为高渗透率样品，问这个样品属于第一、第二和第三种沉积微相的可能性各有多大？

解：由贝叶斯公式可计算取出的一个样品为高渗透率样品(事件 B)时，该样品为第一、第二和第三种沉积微相的概率分别为

$$P(A_1/B) = \frac{50\% \times 1\%}{2\%} = 25\%$$

$$P(A_2/B) = \frac{25\% \times 2\%}{2\%} = 25\%$$

$$P(A_3/B) = \frac{25\% \times 4\%}{2\%} = 50\%$$

故有：当已知取出的一个样品为高渗透率样品时，这个样品属于第一、第二和第三种沉积微相的可能性分别为 25%、25% 和 50%。

2. 逆概率树

贝叶斯公式可用逆概率树清晰地图示表达，例 1－6 中的逆概率问题可用图 1－3 中的逆概率树来表达，为了构建图 1－2 中的逆概率树，还需要计算出已知取出的一个样品为低渗透样品(事件 C)时，这个样品属于各沉积微相的可能性，即

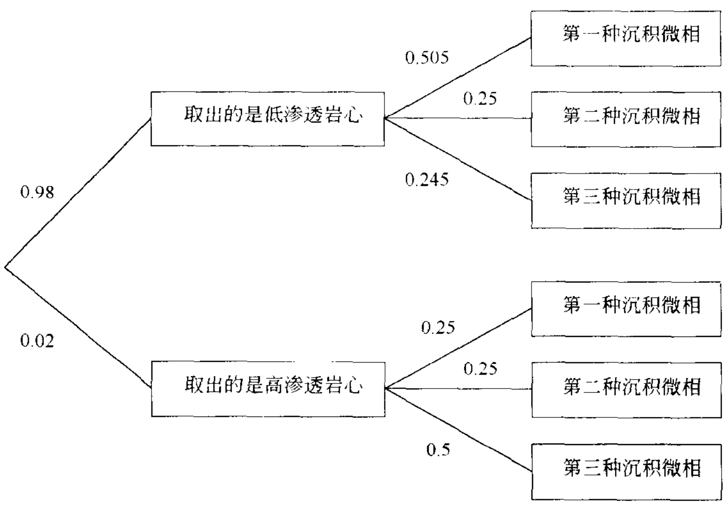


图 1－3 逆概率树