多层次模型分析导论

■ Ita kreft Jan De Leeuw 著 ■ 邱皓政 译 ■ 郭志刚 校

Introducing Multilevel Modeling

重庆大学出版社
http://www.cqup.com.cn

多层次模型分析导论

■ Ita kreft Jan De Leeuw 著 ■ 邱皓政 译 ■ 郭志刚 校

Introducing Multilevel Modeling

重庆大学出版社

版贸渝核字(2006)第 100 号。

图书在版编目(CIP)数据

多层次模型分析导论/(美)克雷伏特(Kreft, I.)，(美)里夫(Leeuw, J.)著；邱皓政译. —重庆：重庆大学出版社，2007.4
(万卷方法)
书名原文：Introducing Multilevel Modeling
ISBN 978-7-5624-4060-4

Ⅰ.多… Ⅱ.①克…②里…③邱… Ⅲ.统计模型—高等学校—教材 Ⅳ.C8

中国版本图书馆 CIP 数据核字(2007)第 041897 号

多层次模型分析导论

Ita Kreft
Jan De Leeuw 著
邱皓政 译
郭志刚 校

责任编辑：雷少波 罗 杉 版式设计：雷少波
责任校对：邹 忌 责任印制：张 策

*

重庆大学出版社出版发行
出版人：张鸽盛
社址：重庆市沙坪坝正街 174 号重庆大学(A 区)内
邮编：400030
电话：(023)65102378 65105781
传真：(023)65103686 65105565
网址：http://www.cqup.com.cn
邮箱：fxk@cqup.com.cn(市场营销部)
全国新华书店经销
自贡新华印刷厂印刷

*

开本：787×1092 1/16 印张：11.75 字数：230千
2007 年 4 月第 1 版 2007 年 4 月第 1 次印刷
印数：1—3 000
ISBN 978-7-5624-4060-4 定价：35.00 元

原序

这一本书是写给社会科学领域中，没有很深的统计与线性代数基础，但是对于传统的线性模型，例如 **ANOVA** 与回归有一定认识的研究者或学生。

本书主要是以随机系数模型来处理多层次模型，这种模型具有固定变量与随机系数。第 1 章介绍了这个模型的基本概念与专有名词，并利用一些研究范例来说明其使用时机。第 4 章则以 **NELS88**（**National Education Longitudinal Study of 1988**；详细的编码请见附录）的实际数据来进行分析与解释的示范。对于这个数据库，读者可以从下列网站得到（**ftp://ftp. stat. ucla. edu/pub/faculty/deleeuw/sagebook**）[1]。本书所使用的软件是由英国伦敦大学教育学院研究人员所发展的 **MLn**，可以处理多层次的数据。本书的实作部分，我们会以 **MLn** 的符号和语法来转换各种不同模型的方程式。在书中，**MLn** 的语法是放在方框当中，随着模型的变化，列出语法的变化，这些语法非常容易理解与阅读，读者并不需要花费太多力气去记忆指令和符号，我们会适时地加以说明。本书对于 **MLn** 的使用者来说非常的实用，但是本书的内容并不只限于 **MLn** 的使用者阅读。

本书与其他著作的不同在于，它是从研究者的角度出发，书中所引用的例子都是可能实际用来检验的模型，透过报表，可以看到一些重要的数据与信息。从实际操作的角度来看，本书特别着重于问题的澄清与方法学素养的提升。对于随机系数模型，我们也是采取保留与审慎的态度，尤其是传统模型与多层次模型的适用性与使用时机问题，我们认为其中有着一些两权消长、各有优劣的空间。本书所提到的阶层性嵌套数据（**hierarchially nested data**）的一些分析流程与替代做法，都是基于我们在教学与研究过程当中的经验，以及我们在多层次模型的网络社群当中所讨论的心得。如果读者有兴趣参与讨论社群，可以寄发电子邮件到 **mailbase@mailbase. ac. uk**，主旨为“**join multilevel your name**”。

最后，我们要感谢 **UCLA** 的 **Mahtash Esfandiary** 与荷兰莱顿大学 **University of Leiden** 的 **Rien van der Leeden** 对于本书部分章节的审阅，他们的意见提高了本书的可读性。

作者的提醒

由于本书的目的在于介绍统计方法，读者必须了解背后的一些限制，只有在一

---

[1] 该链接已不复提供数据文件，经与作者联系，新链接为http://gifi. stat. ucla. edu/sagebook，读者可自行下载数据库，转换成所熟悉的数据格式来进行演练。——译者注

译序

Kreft 与 **de Leeuw** 的这本书，是一本十分有趣且内容丰富的小书。今年初过春节的时候，闲来把它翻了一遍，发现他们能够在不到 140 页中，把多层次模型分析的概念、操作与解释，乃至于一些重要的议题，不拖泥带水地交代清楚，十分佩服，因此决意把它译成中文书，让国内研究者与学生可以很快地进入多层次模型的世界。

唯一遗憾的是，我自己以及诸位同事在分析多层次模型时，多使用美国 **SSI**（**Scientific Software International**）的 **HLM6** 软件，但是 **Kreft** 与 **de Leeuw** 是以英国伦敦大学发展的 **MLn** 为主，因此在阅读范例与数据时觉得格格不入。为了兼顾翻译的信达雅，以及个人的偏好，并考虑市场上 **HLM6** 的高度占有率以及未来的普及趋势，原文中有关 **MLn** 的介绍、运用与结果虽都完全保留下来，但是我另外以 **HLM6** 软件重新把作者所提出的绝大部分模型进行分析，分析的步骤与结果，列在第 2、3、4 章的最后，有部分模型并不是多层次模型，我则以 **SPSS13** 来处理。有兴趣的读者可以按图索骥，利用本书所附的数据库进行演练，将会事半功倍。此外，本书列举了很多网络链接与文献，如果想要深入了解阶层线性模式（**hierarchical linear modeling**）或其他多层次模型的读者，可以自行搜寻有关的信息，相信会有丰硕的收获。

本书的翻译并不困难，最大的收获是我个人从中的学习与体验。虽然我很早就处理过多层次数据分析的问题，也在课堂上教授这些高等统计的应用，但是总是点到为止，并没有机会好好深入了解这门学问。或许可以归咎于分析工具普及与便利性不高，但是最大的障碍是“隔行如隔山”的学科隔阂，如果不是身在量化方法与计量领域的有利位置，我还真没有机会一探多层次模型分析的究竟，尤其是要把这些东西写给别人来看时，要求彻底了解的压力就更明确了，有趣的是，翻译这本书，原本是想把这种新兴的技术介绍给更多的朋友，到最后受惠最大的却是我自己。

如同过去一样，这本书的完成，还是要感谢家人的支持与身边一些朋友的鼎力相助，像温福星老师与林碧芳老师费心的校阅修订，他们的协助让本书的可看性与正确性提高了不少。

译完这本书的最大心得是，我们真的是活在一个知识爆炸的时代。科技发展的脚步实在惊人，新知识、新科技、新想法不断提出，令人目不暇接，如果自己一直停留在自己舒适的空间，真是会有生于忧患死于安乐的遗憾。就像我们“台湾统计方法学学会”的好朋友们一起努力探讨结构方程模式（**structural equation model-**

ing,SEM)的正确运用的同时,阶层线性模式又在学术圈快速发展。很感动的是学会的伙伴们十分能够理解这种发展趋势,大家在 SEM 的探究之外,又"分案"开启了另一个有关 HLM 的探索空间。除了我以外,近期内还会有其他的 HLM 相关著作推出,例如东吴大学国贸系的温福星教授将会有一本 HLM 的专著出版。我们看好 HLM 的一个主要原因,在于 HLM 的方法论与技术取向在社会科学研究中占有相当重要的地位,甚至可以用"相见恨晚"来描述我们的心情。

为什么说相见恨晚,因为我们早就应该使用 HLM 技术来处理多层次的数据分析了。社会科学的量化研究,除了实验方法之外,多是以问卷、量表搜集众人的意见与经验,样本的取得很难做到随机,因此一群群的受测者,就可能因为具有组内的同质性,必须利用阶层化分析技术来处理组内相关(ICC)的问题。我实在很难想象,如果不用多层次模型分析就会得到扭曲的结果的话,那么我们过去几十年来所从事的各种组织、教育、社会、心理学等社会科学研究,究竟产生了什么知识与发现。这早已超越型一错误或型二错误的决策观,这实实在在是一个严肃的方法学问题,无法回避,只能面对。

相见恨晚不要紧,怕的是再次擦身而过,那就会是永远的遗憾了。

邱皓政

2006 年春谨识于

辅仁大学心理学系

心理计量实验室

原序

这一本书是写给社会科学领域中，没有很深的统计与线性代数基础，但是对于传统的线性模型，例如 **ANOVA** 与回归有一定认识的研究者或学生。

本书主要是以随机系数模型来处理多层次模型，这种模型具有固定变量与随机系数。第 1 章介绍了这个模型的基本概念与专有名词，并利用一些研究范例来说明其使用时机。第 4 章则以 **NELS88**（**National Education Longitudinal Study of** 1988；详细的编码请见附录）的实际数据来进行分析与解释的示范。对于这个数据库，读者可以从下列网站得到（**ftp://ftp. stat. ucla. edu/pub/faculty/deleeuw/sagebook**）[1]。本书所使用的软件是由英国伦敦大学教育学院研究人员所发展的 **MLn**，可以处理多层次的数据。本书的实作部分，我们会以 **MLn** 的符号和语法来转换各种不同模型的方程式。在书中，**MLn** 的语法是放在方框当中，随着模型的变化，列出语法的变化，这些语法非常容易理解与阅读，读者并不需要花费太多力气去记忆指令和符号，我们会适时地加以说明。本书对于 **MLn** 的使用者来说非常的实用，但是本书的内容并不只限于 **MLn** 的使用者阅读。

本书与其他著作的不同在于，它是从研究者的角度出发，书中所引用的例子都是可能实际用来检验的模型，透过报表，可以看到一些重要的数据与信息。从实际操作的角度来看，本书特别着重于问题的澄清与方法学素养的提升。对于随机系数模型，我们也是采取保留与审慎的态度，尤其是传统模型与多层次模型的适用性与使用时机问题，我们认为其中有着一些两权消长、各有优劣的空间。本书所提到的阶层性嵌套数据（**hierarchially nested data**）的一些分析流程与替代做法，都是基于我们在教学与研究过程当中的经验，以及我们在多层次模型的网络社群当中所讨论的心得。如果读者有兴趣参与讨论社群，可以寄发电子邮件到 **mailbase@mailbase. ac. uk**，主旨为 "**join multilevel your name**"。

最后，我们要感谢 **UCLA** 的 **Mahtash Esfandiary** 与荷兰莱顿大学 **University of Leiden** 的 **Rien van der Leeden** 对于本书部分章节的审阅，他们的意见提高了本书的可读性。

作者的提醒

由于本书的目的在于介绍统计方法，读者必须了解背后的一些限制，只有在一

---

[1] 该链接已不复提供数据文件，经与作者联系，新链接为http://gifi. stat. ucla. edu/sagebook，读者可自行下载数据库，转换成所熟悉的数据格式来进行演练。——译者注

些前提条件被满足的情况下，本书的目的才可能达成。

第一个条件是，读者必须了解此一方法跟其他方法一样，所能够提供的答案跟数据的本质与搜集的方法有关，对于复杂的人类世界而言，统计方法不可能是完美的工具。

第二个条件是，读者必须了解多层次分析是另一种用来找到数据背后的结构，进而能瞥见真相的策略。就像其他的策略一样，有时候可能会不管用。

第三个条件是读者必须了解本书所介绍的多层次模型是基于某些假设成立的情况，如果这些假设不成立，或许还是可以进行多层次分析，但是标准误、显著性检验等一些统计程序都可能会有问题。当然诸如斜率为随机的这种假设，可以通过样本数据来进行确认，但是如果研究结论是斜率为非随机，并不意味着这个假设一定是错误的，很可能只是在某一个样本上得到这种结果。可能真相是支持这个假设的，只是统计方法无法证明它而已。

第四个条件是，多层次模型的使用者必须了解统计模型是一种数学的模型，如果统计模型所建构出来的数据近似于作者所假设的状况，那么研究结论符合真相的机会会较大。事实上，由于研究者所探讨的问题的复杂性，统计模型的提出往往偏离真相。复杂的模型虽然可以模拟真相，但是由于过于复杂，却削弱了统计方法的效用。如何整理、分析复杂的数据不是一蹴而就的工作，复杂的统计模型尤其难以解释，而且不容易重复观测到相同的结论。复杂的模型先天上就会对于微小的变化非常敏感，造成参数估计因为小小的变动而有很大的不稳定性。

第五个也是最后一个条件是，多层次模型的使用者必须了解这种统计工具之所以有效，是因为数据具有多层次的结构，或是因为理论背景支持使用这种分析，或是因为我们对于数据本身有相当的了解。因此，在进行多层次分析之前，有必要对于数据进行先期的探索检验。

多层次模型这种复杂的统计模型或许是符合真实现象的模型，但是在数据探索的阶段我们并不建议使用，我们也不建议使用过于复杂的模型，例如带有许多解释变量的模型，或是带有复杂的跨层级交互作用的模型。

本书中，都是以小模型来进行示范，变量的选择都具有理论基础，对于数据的性质也有充分的掌握。

目　录

概 说 1

1.1 绪 论

本章的目的在于简介多层次模型的发展历史、使用时机，以及软件的应用。有关应用的部分，本章将从几个不同领域的研究实例来说明。针对阶层性嵌套数据分析的一些概念与名词也将在本章介绍。最后，则对多层次模型的历史进行简要回顾，并以几种软件的介绍作为结语。

1.1.1 阶层、宏观层次与微观层次

阶层数据结构(hierarchical data structure)在社会与行为科学领域中非常常见。如果个体身处不同的组(团体)，此时有些变量与个体有关，有些变量则与团体有关。例如在学校所搜集的学生资料，可能包含一些用来描述学生特征的变量，如社经地位、对于写作业的态度、性别、种族背景；还有一些变量则在描述学校，例如公私立之别或学校类型。学校效能(school effectiveness)的研究者在进行他们的研究时，同时会搜集这个层次的数据，藉以了解学生微观层次与学校宏观层次对于学习成绩的影响。这种在学校所从事的研究的例子以及其他许多类似的状况，说明了我们需要一种可以同时处理多层次测量数据的各层次数据的技术。

多层次模型被设计用来分析阶层结构的数据。在进入详细的介绍之前，我们先将“阶层”一词加以介绍。所谓阶层(hierarchy)系由较低层次观察数据嵌套(nested)[1]在较高层次之内的这种数据结构所组成。例如学生嵌套在学校之内、员工嵌套在公司之内、重复测量嵌套在个体之内。最低层次的测量称为微观层次(micro level)，其他高层次的测量则属宏观层次(macro level)。宏观层次通常系由不同的组别(groups)构成，更正式的说法是不同的脉络(contexts)。因此所谓脉络

---

1 nested 除了被译成嵌套以外，亦可被翻译成内属、嵌套、内嵌等。在本书当中，这几个翻译可互换使用。——译者注

模型(contextual models)一词是指兼具微观与宏观层次数据的模型。脉络模型有时仅有两个层次,例如学生(微观层次)嵌套于学校(宏观层次)之内,亦有可能超过两个层次,例如学生嵌套于班级,而班级则嵌套在学校之内。当然研究者可以想到更多的层级,例如学生嵌套于班级、学校、州别、国家地区等。一旦能够理解阶层的特性,我们可以发现它几乎无所不在。

1.1.2　多层次模型

如果一个模型包含了不同层次的测量变量,称为多层次模型(multilevel model)。在多层次模型中,各脉络可估计出一条低层次的个别直线方程式。在这条回归线中,通常各脉络都有相同的解释变量(explanatory variables)[2]与结果(outcome)[3],但是有不同的回归系数。这些个别回归方程式被一个高层次模型所联结,在高层次模型中,第一层次的回归系数可被第二层次的解释变量所解释。

用来联结这些个别回归方程式的高阶模型的特性,决定了整个模型对于数据分析的方式。在实务上有多种处理方式,最开始的起点是没有高层模式来整合个别方程式的模式,此时每一个脉络各拥有一个回归方程式。虽然这是很自然应该为之的做法,但是从统计的观点来看,此一做法并没有任何新意。

进一步讲,将第一阶层各脉络的回归系数作为第二阶层的反应变量(response variable),称为"斜率结果(slope-as-outcome)"分析(Burstein, et al., 1978)。从统计的观点来看,在各组内与各组间所进行的回归分析彼此并无关联,他们仍是个别进行的分析。不论是未经联结或已被高层模式联结的分析,回归系数都是固定数值(fixed),而非随机变动(random)。如果一个模型利用了全部数据进行分析,则被称为变动系数模型(varying coefficient model),此种模型的分析方法即如其名,是把每一个组分开进行估计,因此每一组拥有各自的一组回归系数。

每一个组进行个别回归分析之后,再以高层解释变量来解释第一层系数的这种说法,尚不足以说明多层次模型的内涵。基本上,多层次模型的基本特性是将研究者所关心的特定层次的数据,用不同的模型通过统计的整合来加以分析。

最简单的整合模型是随机系数模型(random coefficient model)。这种模型中,第一层的回归系数在第二层被以随机变量来处理,这意味着第一层回归系数是从某个机率分配取样而得,此一分配最重要的参数——平均数与方差,可从多层次模

2　解释变量即为自变量(independent variable),在回归分析中又称为预测变量(predictor or predictive variable)。——译者注

3　结果即为因变量(dependent variable),在本书多以反应变量(response variable)称呼之,有时则称为结果变量(outcome variable)。——译者注

型中估计得出的参数所获得。一般来说,在随机系数模型中加入第二阶层的解释变量是一般通则性的做法,也是非常有用的策略,因此一般被通称为多层次模型。

在本书中,将组(groups)与脉络(contexts)视为同义词而交互使用,是指其为阶层数据当中第二层(或较高层)的分析单位。为避免混淆,本书所使用的组别概念有别于实验心理学所惯用的实验组与对照组的组别概念。本书所谈的组别是自然形成的归类结果,例如学校或公司。而脉络则是指与组别相同的概念,而非涉及社会学所谈的更广义的概念。在以下的章节中,我们将以实际的例子来说明阶层性嵌套数据的形貌。

1.2 范 例

在本节当中,我们将介绍一些在不同领域涉及阶层嵌套数据的实际范例。第一个例子是企业员工嵌套在公司之内。研究发现指出,在不同阶层所进行的不同层次分析所得到的结果不尽相同。在这个例子中,我们会讨论到各脉络的观察值具有相依性。第二个例子则是学生嵌套于学校的班级之内。其他的例子则为不同领域的实际范例。

1.2.1 企业员工的薪资水平

第一个例子是 Kreft 等人(1995)所进行的研究。研究资料是从 12 个不同的公司所获得,微观层次的解释变量是教育水平,反应变量为薪资收入。公司的类型,例如公营与民营,为第二个阶层解释变量。对于员工层次的分析发现,教育水平与收入具有正向关系,教育水平较高者,薪资收入较高。以 12 个公司为观察对象的公司层次分析发现相反的结果:教育水平与薪资收入呈现负向关系,当一个公司的平均教育水平越高时,该公司员工的平均薪资收入越低。

在此例中,公司层次变量为聚合测量(aggregated measurement),也就是以公司的平均教育水平为解释变量,反应变量是公司的平均薪资。分析结果显示,不同层次的分析得到不同的结果。高层次聚合数据的分析结果与原先微观层次的结果不同的这种现象,在 Robinson(1950)的研究中即已发现,此种聚合偏误现象称为 Robinson 效果(Robinson effect)。前面的组织研究数据中,微观层次的分析发现教育水平对于收入有正向效果,但宏观层次的分析发现负向效果,逻辑上来说,两个层次的教育水平变量是在测量不同的东西,取决于分析的单位。此一结果显示我们有必要将两个层次的分析同时加以处理,因为两个层次的结果都非常重要,而且彼此具有关联。

此一范例显示出阶层嵌套数据的一个重要特性，亦即同一个公司的员工比起其他公司的员工更为相像。公司内员工的同质性可以通过公司的特征来反应，更具体来说可由组内相关（intra-class correlation）（注1）来衡量。如果组内相关很高，团体内具有同质性，而且可能与其他团体的差异很大。此一现象也出现在本范例的结果中，亦即教育水平对于员工收入的解释，在公营企业远高于私人企业。一般而言，如果组内相关较低时，团体间仅有些微差异。如果组内相关低至近乎0，则研究者所关心的变量无组间差异，换言之，同一组内的个体间差异，就像该组个体与其他组个体之间具有相同的差异程度。一个数值为0的组内相关意味着数据的丛集不影响研究变量关系，因此可以忽略组间的差异。如果组内相关存在，数据的嵌套结构就必须加以考虑。忽略组内相关会导致结果的信度问题，但是组内相关必须具有统计的显著性，并且有相当的强度，有关此点可参见 Cochran（1977）的详细讨论。

1.2.2　药物滥用预防研究

第二个例子为药物滥用防治研究（Kreft，1994）。在该药物防治研究中，研究者所关心的是高中青少年药物防治计划是否有效。实验处理是防治计划，测量与分析的对象是学生。变量则依不同的阶层，在学校/班级层次与学生层次中有所不同。参与研究的学校是随机抽样得来的，学生所属的班级也可以被视为是从某一类特性的学校所随机选择的样本。研究所测量的变量包括学生的风险因子，例如心理因素、学业成就、贫富差异等，在学校或班级层次的风险因子则包括一个学校药物滥用的程度、邻近区域药物滥用的状况。

有关药物防治的文献指出，个体风险因子与药物防治计划之间具有交互作用，而学校风险因子也与防治计划具有交互作用。文献中还可以看到许多有关脉络与学生特性的假设效果。若以多层次分析术语来说，这些关系应称之为跨层级交互作用（cross-level interaction），因为变量的关系横跨学校与学生层次。可以预期的是，某些学生（例如高危险群学生）容易受到某种环境的刺激而使用药物，但是在其他的环境下则可能减少他们使用药物的机会。为了检验这些研究假设，我们所需要的不仅是嵌套的数据结构，而且要能够估计跨层级的交互作用。

1.2.3　学校效能研究

第三个例子是学校中老师的教学效能研究，在这个例子中，阶层结构扮演重要的角色。研究所关心的对象是学校与老师，同时也包含了学生。研究者可能对于学生个人以外的因素，例如组织结构如何影响学生的学习成果感到兴趣，研究者也可能关心老师的教学经验、智能、教学风格等因素是否影响学生的学习。典型的例子是 Cronbach 与 Wenn（1975）、Burstein 等人（1978）与 Aitkin 和 Longford（1986）所