分析单位、分层结构、分层模型

郭志刚
北京大学社会学系

摘要：分析单位在研究中具有重要意义，直接关系到研究结果的有效性。值得注意的是一些错误使用分析单位的做法在实际中还很普遍，本文分析了若干例子，并指出它们的问题所在。本文还讨论了与分析单位相连的一些方法论谬误，指出它们的问题出在类比推理的可错性上，并讨论了在假说检验中的论证和验证。本文还介绍了分层结构的概念，说明分层结构现象以及有关数据在很多研究领域都是普遍存在的，然而由于常规统计方法所存在的种种局限，很难将不同层次分析单位数据真正结合在一起综合分析。近年来分层模型的发展为分层结构数据的分析提供了必要的手段。本文简单介绍了分层模型的基本原理，并就分层模型的不同子模型介绍了这种分析方法的功能。分层模型不仅提供了一种现实的统计分析框架，而且对于其他有分层特征的研究对象的理论探讨也具有参考价值。

1. 分析单位

分析单位是研究方法论中的基本概念之一（艾尔·巴比，2000：119），特别是对统计分析等操作化要求较高的研究而言，具有非常重要的意义。实际上，分析单位是对研究对象的具体界定，而研究对象是研究关注点的载体，在统计分析中称为案例。在社会学研究中，关注点一般关系到研究对象的特征、取向或行为，这些关注点构成了研究变量，用以对研究对象进行分类、筛选、比较、或分析变量之间的关系。在社会学研究中，分析单位既可以是个人，也可以是群体、组织、社会制品或事件等。所有满足一定条件的研究对象的集合构成了研究的总体。比如，人口便是某一范围中所有个人构成的总体[1]。对某一总体中个体所承载变量进行研究，并不一定要求普查所有个体。在现代科学研究中，经常采用抽样调查的方法来收集一部分个体样本的信息[2]，然后通过样本数据分析对总体情况做出推断估计和假设检验。

确定研究的分析单位是研究设计的重要工作，分析单位确定得好不好关系到研究结果的有效性。比如夫妇年龄差是婚姻研究的一个重要指标，对于一个人口中所有夫妇的年龄差常常用平均的夫妇年龄差表示。但是如何根据人口普查原始数据来计算这个平均夫妇年龄差呢？一种常见的、比较省事的做法是，先选择人口中所有的有配偶男性并计算出其平均年龄，再用同样方法计算出有配偶女性的平均年龄，最后再计算两种平均年龄之差来作为平均夫妇年龄差。但是严格地说，这个差并不是夫妇年龄差。因为事实上，夫妇年龄差的分析单位不是处于有配偶状态的个人，而是婚姻（涉及两个人的群体）。因此正确的做法应该是，先用普查原始数据的有关信息对每户中的夫妇进行匹配，匹配出的夫妇双方的信息被另存为一个案例，最后形成一个新的以夫妇（即婚姻）为分析单位的总体[3]。每一对夫妇都能计算其年龄差，形成一个变量，这个变量的平均数就是所有夫妇的平均的年龄差。至于前一种省事的方法所存在的主要问题是：1、由于绕过了匹配夫妇这一环节，因而根本没有以夫妇为分析单位来计算。而所得的结果有时可能与严格口径得到的夫妇年龄差存在很大差别[4]。比如，当社会中存在许多老夫少妻和老妻少夫情况时，上述省事方法却不能反映这些情况对夫妇年龄差的影响。于是，所计算的平均年龄之差就会大大低于真实的平均夫妇年龄差。2、当有配偶人口中存在许多配偶不在普查范围中的情况时，这些人的配偶的年龄及夫妇年龄差实际上根本得不到，因此不符合研究对象的条件。而上述省事方法并未将这些人从计算中排除出去。3、这种省事方法只是在汇总层次计算了一个差，并没有真正形成以婚姻为分析单位的夫妇年龄差变量，因而不仅这个差根本不是夫妇年龄差的平均值，而且也不能提供任何有关夫妇年龄差分布的测量，如按夫妇各自

[1] 很有意思的是，在英文中人口与总体是一个词，population。

[2] 有时是为了节约研究成本，有时则因为总体是无限的（如大气、海水），因而根本无法普查。

[3] 人口普查是以户为调查单位实施的，因而户是调查单位。而调查中每户都收集所有户成员的信息。利用人口普查数据在户内匹配夫妇，以得到婚姻为单位的数据是可行的，具体做法参见郭志刚（1999a:14）。

[4] 弗庭格尔（1995:107）在介绍家庭生命周期统计测量时曾指出过这种方法的错误。

年龄的交互频数表或不同年龄差的分布比例或年龄差的方差，而这些指标同样是婚姻研究的重要信息[5]。

有时不严格地定义分析单位，便不可能深入地分析研究对象。比如对早婚的研究，既可以从个人角度来分析，也可以从婚姻的角度来分析。判断个人是否早婚只需要看个人的初婚年龄；但是判断一个婚姻是否属于早婚则需要看夫妇双方的初婚年龄，并且只要有一方早婚这个婚姻便应该定义为早婚。在一个早婚研究中（郭志刚，1999b:6），以个人为分析单位时可以发现北京市男性早婚的数量是女性早婚数量的 3 倍，而进一步以婚姻为单位的分析便揭示出，男性早婚者当中有 80%以上是因为女方比男方年龄大，结果是男方属于早婚而女方并不是早婚，甚至有 20%的早婚婚姻中女方结婚时已超过晚婚年龄。这种情况说明，北京绝大多数的早婚婚姻都是由于女方的年龄确实比较大而不能再推迟结婚的实际情况所引起的。然而，以个人为分析单位虽然能够发现男性早婚多的现象，却不可能找到这种现象背后的主要原因。

综上所述，确定适当的分析单位是研究中很重要的一环。

2. 与分析单位相关的方法论谬误

对于同样的研究关注点，可以在不同的层面上开展分析。比如，研究供给与需求之间的关系，既可以从全国的角度（汇总统计指标）来建立它们之间的函数，也可以从厂商的角度来加以分析。这又涉及到选择分析单位的问题。依据分析单位的口径，我们将研究分成宏观分析和微观分析。宏观分析和微观分析在理论上的不同，背后也同时反映了研究方法论上的不同。显然，这两种研究都是必要的，它们既有各自的独特贡献，又是相互补充的。但是这里所说的互补，并不是指不同层次的结论必须简单化的一致，而是指微观运行机制和宏观运行机制可以整合在一个综合框架中而并行不悖。

事实上，事物的宏观运行机制和微观运行机制经常是不一致的。萨缪尔森在讲经济学的研究方法时说："在经济学领域中，十分肯定的是，对于个人说来是对的东西，对整个社会说来并不总都是对的；反之，对大家说来是对的东西，对任何个人说来可能是十分错误的。"（转引自王治平，2002:20）比如，某个农民特别辛勤地劳作，秋后可因其产量多而增加收入，但若全体农民都努力干活且老天又风调雨顺，农业获得全面丰收，此时农民的总收入却要下降，因为谷多而贱，谷贱伤农（王治平，2002:21）。

既然事物的宏观机制和微观机制可能是不一样的，所以根据某一层次观察材料做出的分析结果是不能简单推论到另外的层次去的。也就是说，当分析单位不同时，不能进行跨跃层次的推论[6]。根据研究方法论的原则，将微观单位数据的分析结果推论到宏观单位时便犯了"生态学谬误"(ecological fallacy)；而将宏观单位的分析结果推论到微观单位时便犯了"简化论"(reductionism)谬误[7]。

为了避免误解，先澄清一下有关推论和谬误两个概念。从研究方法论而言，推论是产生新思维的一种手段，其背后的逻辑是类比推理。然而类比是一种不严格的推理，有可错性，也就是说推论既可能是对的，也可能是错的。借助这种不严密的推论来产生新的假说是可以的，但是作为对这种假说的证明却被视为方法论上的谬误。按本文作者的理解，方法论中所说的谬误(fallacy)与通常所说的错误(mistake)有所区别，其不同点在于它们分别用于表示两种不同性质的检验得出的否定结果。艾尔•巴比认为："科学的两大支柱就是逻辑和观察。科学对世界的理解必须（1）言之成理，并（2）符合我们的观察。"（2000：35）也就是说，科学理论要经得起逻辑上的检验，我们可称之为论证；同时，科学理论也要经得起实际的检验，我们可以称之为验证。验证是终极标准。但是验证往往需要具备很多客观条件才能完成，因而本身是成本很高、难度很大的研究工作。相对而言，逻辑上的检验比较容易，算是初级标准。科学理论往往是从假说成长起来的。而科学假说应该已经具备逻辑上的严密性和完整性，只是尚未经过实际检验。在科学界认为有必要对一个假说进行实际检验之前，往往先进行逻辑检验。当方法论上说到谬误时，往往是指推理逻辑有问题，即涉及的是论证方面的否定。而在说到错误时，则更多的是指观点不符合事实，即涉及的是验证上的否定。所以，借助类比推论以形成新思维或假说时是可以的，但是要明确，在这个过程中既不存在论证，也不存

[5] 郭志刚、邓国胜（1998:5-6）发现处于严重婚姻拥挤的年龄组会通过扩大夫妇年龄差来克服寻偶困难，但很有意思的是这些年龄组的平均夫妇年龄差与其他年龄组没有显著差别，而这些年龄组的夫妇年龄差的方差却有显著不同。这说明，扩大择偶年龄范围是双向的，既包括年龄更小的，也包括年龄更大的。因此，平均数变化不大，而方差变化则较明显。

[6] 这里所说的与从样本向总体的统计推断是两码事。实际上，样本和总体的分析单位是完全相同的。并且，样本向总体的推断仍然有错误的可能性。

7 在各学科中所用名称有所不同，比如将微观结果推论到宏观，在社会学中常称为简化论，在经济学中则称为合成谬误或构成谬误(fallacy of composition)，还有的学科将其称之为原子论谬误(atomistic fallacy)。而国内有关方法论文献中对这些概念的译法也很不统一。本文中所用译法遵循尽量贴近英文原词义和有关概念原义。

在验证。将建立假说时已经用过的经验材料又反过来再用作该假说的检验证据，是又一种方法论谬误，称为事后检验假设(ex post facto hypothesizing)。

实际上对生态学谬误和简化论的方法论意义上否定，并不在于这两者做法都使用了推论，而在于它们将可错的、不严密的推论误解为是不会错的、严密的推论。其实，不严密的推论至多只是得到一种新的假说而已，要成为理论（已被证明是正确的假说，更严格地说是已经经过若干事实检验但尚未被证伪的假说）不仅在推理上已经存在漏洞，在逻辑上还不如从根据本层次经验事实结合有关命题推出的假说，更不要提作为假说还须通过事实的检验。其实，将不严密推论当成了严密推论（不会错）时，本身就是一种推理谬误（尽管其推论也许事实上真的是正确的）。这两种谬误的共同点是，在完全不讨论前提条件的情况下，将一个层次的经验结论与另一个层次的经验结论划了等号。它们的区别只是推论从层次的方向上是相反的。这两种谬误的来源，实际上是将某一学科中的所强调的因素无限推广。比如，生态学强调大环境对寓于其中的物种的决定作用，持“大决定小”观点。而简化论则认为事物的性质是由其最基本要素的性质决定的，持“小决定大”的观点。但是，这里我们得有一点辩证思想，应该说宏观影响和微观影响的确都存在，需要具体情况具体分析，将任一一个加以绝对化，来排斥另一个的存在，都显然是不对的。

社会学研究方法的教科书在分析生态学谬误时往往都引用一个典型的例子，来说明不能根据以地区为单位得到的某一类人所占比例与犯罪率之间的相关做出这类人犯罪多的推论。其实说得更透彻一点，在犯罪研究中，直接的分析单位应该是个人，而不应该是地区。因为犯罪是个人行为，犯罪的个人是犯罪和某些属性的共同载体，在他们身上体现着犯罪与某些个人特征之间的联系。而地区是不可能犯罪的，地区犯罪数据及其他人口特征数据是从个人数据汇总出来的，而这种汇总过程已经割断了犯罪与其他属性之间的直接联系。因此，地区汇总数据之间的相关不能推论到个人[8]。

但是，我们在研究犯罪问题的时候，确实又不能忽略大环境的影响。一个地区的人口构成、制度、风气也会影响个人行为，这种生态的、或宏观的影响的的确确是存在的，有时甚至是极为重要的。应该说，这种情况并不局限于犯罪研究，而是很多研究所面临的普遍性问题。

在这种情况下，如何将宏观分析和微观分析整合在一起便出现了操作上的困难。虽然现在很多研究都自称要将宏观分析和微观分析相结合，但实际所做的只是简单地将不同层次的分析结果比较一下而已。这种“结合”并不太难，但真的要将两者整合起来无论是在理论上还是在方法上都是一件极为困难的工作，因为我们面临的研究实际上需要在框架中**同时**包括不同层次的分析单位。

3. 常规统计方法面对分层数据的困窘

面对复杂的社会事物，仅用宏观或微观来概括分析单位的性质显然是太简单了。实际上，从微观到宏观存在着多层的分析单位序列。比如：个人、家庭、社区、……、国家。它们一级套一级，有种归属关系。从统计上，称之为嵌套关系(nested)，比如个人嵌套于家庭，家庭嵌套于社区，等等。每一层分析单位都有自己的特征和属性。个人有性别、年龄、身高、体重等生理特征，还有其他社会经济特征。较高层次分析单位的特征中，有的是从下级分析单位的特征汇总而来，比如人口数、平均年龄、平均家庭规模等，有的属性则是从另外渠道获得的，比如自治区、经济特区、山区、城乡、企业所有制和单位所属行业等。这种情况称之为分层数据结构。实际上，分层数据结构在不同领域研究中普遍存在。

下面本文就学生学习成绩影响因素的例子来讨论常规统计分析方法的局限是如何阻碍这一研究的，借此反映许多研究领域中的共同问题。

影响学生学习成绩的因素是什么？首先可以从个人角度考虑可能存在的种种因素，如智力、用功程度、家庭社会经济背景。此外，还可以从学校角度考虑学习过程所处的环境，如学校的条件、教师的水平，学习风气等。暂且仅考虑这两个层次，便足以说明问题了。

这里，假定数据来自抽样调查，先抽中若干学校，然后再在抽中的学校中抽取若干学生。于是在数据中，学生是第一层分析单位，学校是第二层分析单位。学生嵌套于学校。经验和理论告诉我们，学生个人

[8] 当分析单位已经不再是因果过程真实发生的层次时，变量之间所表现的关系便有可能是一种伪相关。例如人口学中的一个常识是，现在发达国家的人口死亡率较高。。但是，我们并不能推论个人越富裕死亡概率越高。因为，这一现象背后的真正原因是发达国家的老年人比例很高，因而人口死亡率较高。控制了人口年龄结构影响，发达国家的人口预期寿命远高过发展中国家。

因素和学校环境因素对于学生的学习成绩都很重要。这在理论分析中没有问题，并可以很简单地用因果框图形式将两大方面的因素并列作为学生成绩的原因。但是，如何从量化统计角度检验这两种层面的影响是否真的存在、两种影响到底哪一个更大，问题就困难多了。因为常规统计方法（如功能强大的回归分析）只能处理同一分析单位的数据。要不然以学生为单位来分析（以每个学生为案例，以学生的各种特征为变量），要不然以学校为单位来分析（以每个学校为案例，以学校的各种特征为变量）。

回归分析用回归系数来表达因变量（如学习成绩）与自变量之间的定量关系。求解回归系数的前提条件是，待定回归系数要使得自变量能尽可能多地解释因变量中的差异。换句话说，学生在学习成绩上的变异（用总方差标志）代表了所有影响因素作用的总和，回归系数乘以自变量观测值便得到成绩的估计值（在多个自变量时便将各自变量的单项估计值加起来得到最终估计值），而估计值的方差称为解释方差（又称为回归方差），代表自变量的绝对解释能力。也就是说，通过回归，可以将因变量的总方差分解为解释方差和未解释方差两部分。用解释方差占总方差的百分比，来表示自变量的相对解释能力，称为确定系数。确定系数的值域为（0%，100%），当其达到上限时说明回归方程 100%地解释了因变量的变化（即因变量是回归方程自变量的确定结果，没有任何误差），当其达到下限时，说明回归方程中的自变量完全没有解释能力。其实，回归分析对于总方差的分解，实际上与方差分析（或协方差分析）的内在统计原理完全相同，也就是说，一个完整的回归分析同时包含着方差分析在内。

总而言之，在统计研究中，要想研究某个因变量的影响因素，必须要把这个影响因素作为自变量明确地定义在模型中。要想比较不同影响因素，必须将不同影响因素同时列入模型之中，在控制其他因素的条件下比较。此外，研究自变量之间的互动作用，也必须将代表互动作用的变量（通常用具有互动关系那些自变量的乘积来代表）。互动作用的存在标志着某一自变量对因变量的作用依其他自变量的取值的变化而变化，因而是一种有条件的作用。因此定义模型是研究的关键，定义模型必须反映研究的需要，模型实际上代表了某种理论假设。而总方差的分解则有助于了解模型的解释能力和有待于进一步解释的余地。

回到学习成绩的例子。虽然我们知道，学生个人特征的影响和他们所在学校特征的影响（假如真的存在的话）均已反映在学生的成绩中了，但是常规统计方法的局限性导致在同一模型之中，即在要求同一分析单位情况下，我们没有办法同时考虑个人层次的变量和学校层次的变量。

长期以来，作为没有办法的办法，实际中采用三种不同的做法来处理。虽然这三种做法明显违反了方法论和统计原理，但是往往以“不好的研究要好于不研究”的理由被容忍，一直存在于研究实践当中。其产生的不良后果：一是其方法不严密导致其研究结果可能是错误的；二是久而久之对这些做法会从容忍演变为熟视无睹，甚至还会被当作经典来仿效。

就上述这个例子来说，第一种做法是先将学生层次变量全都汇总到学校层次，所有定距变量都可以取平均值，比如因变量现在变成了学校的平均成绩，而分类自变量汇总到学校一级时便成了分类比例。然后，在学校层次建立模型进行分析。为了尊重原来的研究对象，稍讲究一点的分析还会按学生规模对学校加权。但是这仍然不能弥补这种做法的缺陷。首先，数据汇总统计属于信息概括，而概括的结果总是在突出重点的同时大量精简信息。但问题是，在这种具体场合“精简”则意味着信息损失。学生成绩一旦汇总为学校的平均成绩便抹杀了同校之内的学生之间的差异，这实际占了原来成绩总方差的很大一部分。也就是说，此时模型中因变量上的差异只剩下了学校平均成绩之间的差异。其结果是，汇总变量之间的关系总是显得较为密切，然而这常常与直接分析未曾汇总的变量所反映的情况大相径庭。于是，这种做法不但浪费了信息，使得原生状态的学生成绩差异大打折扣；并且分析结果将只见森林不见树木，如果将以学校为分析单位的分析结果推论到学生，还将产生生态学谬误。如前所述，没有理由推论汇总层次发生的关系同样存在于个人层次，因为所分析的汇总数据中已经切断了本原性因果关系。比如就学生个人来讲，学校的汇总成绩不代表他自己的成绩，学校的其他汇总特征不代表他自己的特征。当某一学校的平均成绩较高同时具有某种其他由学生层次汇总来的显著特征时，并不一定说明该学校的学习成绩好的学生都是具有这种特征的学生。尽管学校提供的学习处境对学生学习过程有重要影响，但归根结底学习成绩是学生所取得的，而不是学校所取得的，因此对学习成绩的研究不能对学生视而不见。

第二种做法是在学生层次进行分析，不再考虑学校层次的变量。显然，这种模型站在简化论的立场上，因为它完全忽略了学生学习过程的处境。该模型背后的理念是，学校的所有特征都与学生成绩无关，而各个学校平均成绩之间的差异应该完全归结于各个学校的学生在个人特征构成方面的不同（因此简化论的另

一个称呼是构成谬误）。经验告诉我们，个人特征变量与其在较高层次上汇总出来的构成变量的意义是不同的，比如一个人本人富有与其居住在富人区是两回事。再比如，“昔孟母，择邻处”，今天的父母也还在请客送礼托人情让孩子进重点校，这样做的用意并不是旨在改变孩子本人的特征，而是在追求外在处境的改善。因此，对学习成绩的研究中完全忽视学校因素对学生成绩的影响也显得过于偏颇。

第三种做法仍是在学生层次分析，但是企图兼顾学生和学校的特征。具体做法是将学校的特征分赋于所嵌套的学生个人案例中去。这样影响学生个人成绩的自变量中，既包括个人特征，也包括学校特征。这种做法虽然看起来有道理，并且在实际中十分流行，但是这种模型的回归求解存在统计方法上的问题。由于学生（个人）是从特定学校（处境）中抽出来的样本，而模型中明确定义的学校变量通常只是学校各种特征中的一部分，因此在应用常规最小二乘法回归求解时，所有模型中未明确定义的其他方面的学校特征而产生的处境差异（本来属于学校层次上的未解释差异）现在便进入了学生层次中每一个学生案例的误差项。这意味着，同一个学校的学生在其误差项中有一个相同的部分，使同一学校学生的误差项之间产生了相关。这种情况违反了常规最小二乘法回归的假定条件，即误差项之间独立，而这一假定是计算标准误和进行统计检验的基础。由于这种处理方法不能保证误差项独立，因此常规回归方法求解的估计有偏，即求出的回归系数值要比真实情况大。此外，这种方法还假定了学校变量对所有学生的影响是一律的，但有时这并不是事实。也就是说，这种模型没有考虑可能存在不同层次变量之间的互动效应，比如学校的制度和风气可能对不同性别、不同智商、不同家庭背景的学生的影响是不同的，而这些很可能正是研究的关注点。

总之，以上三种做法在方法上都存在很大缺陷，不能够很好地解决分层研究框架的问题。并且，有关分析单位的方法论谬误在研究实践中一直能够存在，其原因既有研究人员的方法论素质不够高的问题，也有时是因为数据可得性问题，即所能得到的数据只是单一层次而所要研究的框架却是多层的。但是，在已经拥有多层结构数据的情况下（如上述第三种做法时），问题便出在实际上一直没有适当的统计方法能够处理多层次的研究模型。然而，这种情况已经出现改观，近年来分层模型在技术方法上有了重大发展，这一模型也在科学研究中得到日益的普及。

4. 分层线性模型及其功能

在 1980 年代中后期，分层线性模型(hierarchical linear model)的发展开始趋于成熟(Raudenbush and Bryk, 2002：5-6)[9]。这种模型在不同学科有不同称谓。在教育研究中，它被命名为分层线性模型。在社会学研究中，它经常被称为多层线性模型(multi-level linear models)”；在生物统计学中被称为“混合效应模型(mixed-effects models)”和“随机效应模型(random-effects models)”；技量经济学则称它为“随机系数回归模型(random-coefficient regression models)”；统计学文献则称之为协方差成分模型(covariance components models)。

这种分析方法可以在一个模型之中同时处理微观层次的个人变量和宏观层次的处境变量。分层线性模型在三个方面优于常规统计方法：一是能够对个体单位取得较好的效应估计；二是可以对各层次之间的效应建立模型并进行假设检验；三是可以分解各层次间的方差和协方差成分。这种模型不仅在技术上是强大的，而且比传统单一层次的统计技术具有更大的包容性（或称为一般性）。

下面将简单介绍分层线性模型的原理，从其一般形式出发，分别讨论其在某些特定情况下的几种主要子模型的形式及意义，借以展示分层模型的分析能力。本文并不企望从统计上将这个复杂的模型完全解释清楚，而是希望可以概括出在分层研究框架中，可以容纳哪些方面的研究内容，以及不同子模型所代表的理论假设，并讨论这些研究分析背后的实际意义是什么。其实，这些概念和思路不仅对具体开展分层统计研究十分重要，而且对有关理论研究也有重要的方法论参考价值。

下面还是借助不同学校学生成绩的例子来进行讨论。

假设我们从学校的总体中随机抽取 J 个学校作为样本，其中 J 是一个较大的数量。可以为不同学校写出一个通用公式来表达学习成绩与其他变量之间的关系。

先在学生个人层次建立一个回归模型，表达个人变量 X 与个人成绩 Y 的关系：

微观模型： $$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \overline{X}_{\cdot j}) + r_{ij} \qquad [1]$$

其中，下标 j 表示不同学校，下标 i 表示该学校中不同的学生。β_{0j} 是回归截距，β_{1j} 是回归斜率，r_{ij}

[9] 本文中凡有标号的公式以及符号表示方法都源引自Raudenbush和Bryk的专著《分层线性模型》第2章（2002）。

是个人在应用回归方程$\hat{Y}_{ij}=\beta_{0j}+\beta_{1j}(X_{ij}-\bar{X}_{\cdot j})$估计其成绩时的随机误差。根据回归原理，当用自变量的中心离差（即上式中括号内）表示自变量的变异时，回归方程的截距等于因变量的平均数[10]，而斜率不变。这样一来上式中的截距便是各学校的平均学生成绩，用以表达学校的成绩**水平**；而斜率仍是学生特征上的差异对其成绩的影响效应，用以表达成绩与自变量的**关系**。上述微观模型对应每个学校各有一套回归参数（包括截距和斜率），也就是说共有 J 套回归参数。其实际意义为：

某学校中某学生的成绩＝该校平均成绩水平＋该学生特征影响的成绩增量＋该学生的随机误差。

注意，上式中截距和斜率在一个学校之内对所有学生是共性参数，但其斜率要与学生个人变量结合起来起作用，只有学生的随机误差是纯粹个性的反映。

但是现在并不知道每个学校的平均数和斜率值（即β_{0j}和β_{1j}），还需要根据抽样数据来估计它们。具体地说，可以用本学校的其他特征（诸如资金水平、组织特征、政策）来估计这两个参数。例如，建立一个学校特征的变量 W_j，然后再以学校为分析单位建立两个回归方程来分别估计出这两个参数：

学校层次的截距模型： $\beta_{0j}=\gamma_{00}+\gamma_{01}W_j+u_{0j}$ [2a]

学校层次的斜率模型： $\beta_{1j}=\gamma_{10}+\gamma_{11}W_j+u_{1j}$ [2b]

其中各参数的意义为：

γ_{00}为所有学校平均成绩的平均数；
γ_{01}为某学校特征对该校平均成绩的影响；
γ_{10}为所有学生的特征对其成绩的一般效应；
γ_{11}为某学校特征对其平均成绩的特殊效应；
u_{0j}为各学校在平均成绩的随机误差；
u_{1j}为各学校在斜率关系上的随机误差。

假定 u_{0j} 和 u_{1j} 是平均数等于 0 的随机变量，它们代表着控制学校变量 W_j 以后在学校截距β_{0j}与学校斜率β_{1j}上还存在的差异。现在还不能来估计这些回归方程的各种γ参数，因为其因变量(β_{0j}与β_{1j})尚未得到观测。然而，分层结构数据中已经包含着这一估计工作所需的信息。将方程 2a 和 2b 代入方程 1，就能得到可用于估计[11]的、含有不同层次自变量的组合方程：

$$Y_{ij}=\{\gamma_{00}+\gamma_{01}W_j+u_{0j}\}+\{\gamma_{10}+\gamma_{11}W_j+u_{1j}\}(X_{ij}-\bar{X}_{\cdot j})+r_{ij}$$

从这一通用组合方程可以看到其中的两层分析单位的嵌套关系。上述公式具有 Y＝{A}＋{B}X＋E 的简单回归形式，其中的截距$\beta_{0j}=\{A\}$和斜率$\beta_{1j}=\{B\}$中又各自嵌套着一个内容不同、而形式完整的简单回归式。注意这两个嵌套的简单回归式中，自变量都是宏观单位变量，而误差项 u 也是对应宏观单位的误差。因此，因变量Y_{ij}的变异，从大的方面可分为学校截距变异、学校斜率变异和个人特征变异。通过数据分析，γ类参数是可以估计出来的，而 u 和 r 类随机误差则可以进行方差分析。将随机误差项都集中到回归式的右边后，组合公式变成这样三个部分：

$$Y_{ij}=[\gamma_{00}+\gamma_{01}W_j]+[\gamma_{10}+\gamma_{11}W_j](X_{ij}-\bar{X}_{\cdot j})+[u_{0j}+u_{1j}(X_{ij}-\bar{X}_{\cdot j})+r_{ij}]$$

第一部分$\gamma_{00}+\gamma_{01}W_j$作为学校的**截距估计**。其中的第一项为一般情况（总平均成绩），构成一个总的基准，然后再根据学校具体的W_j情况调整，γ_{01}为调整系数。注意这一部分的所有参数（γ_{00}和γ_{01}）和变量（W_j）完全是来自学校层次的。实际上，这部分的代数和决定某一具体学校学生成绩的回归线的截距，描述的是该学校学生的平均成绩。

第二部分是学校的学生个人变量$(X_{ij}-\bar{X}_{\cdot j})$与其成绩的**关系估计**，即影响作用的斜率。这个斜率的估计由两项成分构成：第一项γ_{10}代表一般情况的平均斜率（实际上为各学校学生特征与成绩的回归斜率的平均数），将其作为一个基准。第二项是根据所在具体学校的特殊情况所进行的斜率调整。注意这个调整

---

[10] 注意这里说的平均数是泛指的。其平均数的类型由计算离差时所用的具体中心而定。在应用原自变量X进行回归时，β_{0j}是X=0的那些学生的成绩（有时其实际意义很难理解）。而在对变量取中改造之后，Y轴会移向X实际取值的中心。在上述公式中X是以所属学校的学生变量的平均数$\bar{X}_{\cdot j}$为中心的，因此截距β_{0j}为该学校的平均成绩。要是以所有学校中的所有学生的自变量总平均数$\bar{X}_{\cdot\cdot}$为中心时，那么截距β_{0j}是所属学校在控制学生自变量影响后的调整的平均成绩。

[11] 尽管这一模型不能采用常规回归分析，但可以采用最大似然估计。需要指出的是，如果u_{0j}和u_{1j}值对每个学校都不起实际作用的话（即设其值为0，因而相应效应可从模型中省略），那么这一组合方程便实际上与常规回归模型等价了。

量实际上体现了学校变量和学生变量的互动产生的效应，$W_j(X_{ij}-\bar{X}_{\cdot j})$ 的积在统计上可视为层际互动变量，而 γ_{11} 则为调整系数。注意，对于同一个学校的学生来说，学校特征变量值 W_j 是一样的。当其调整系数 γ_{11} 不等于 0 时，这个交互效应便开始发生效应，使某一学校中学生特征与成绩之间的关系相对于平均斜率发生偏离。

前两个部分所表达的都是因变量 Y_{ij} 的变异中能够由回归方程“确定”或“解释”的部分。其中的各种回归系数 γ 的值都可以根据数据估计出来，并可以进行推断统计和显著性检验。

而第三部分则是各种随机波动，即因变量变异中尚未被回归方程所解释的部分。从方程的随机波动部分可以看出，这个分层组合方程并不是标准的常规最小二乘法(ordinary least squares, OLS)所假定的典型线性模型。用常规回归取得有效估计和准确假设检验的条件是，随机误差必须独立，服从正态分布，并且有相同的方差。此外常规回归中只有一个误差项，不能再被分解。但是分层模型的组合方程的随机误差则与此相反，具有更复杂的形式 $u_{0j}+u_{1j}(X_{ij}-\bar{X}_{\cdot j})+r_{ij}$。这样的误差在每个学校之内是相关的，因为在某一学校 j 中的每个学生的成分 u_{0j} 和 u_{1j} 是共同的。并且各学校的误差可以有不同的方差，因为 u_{0j} 和 u_{1j} 在各学校取值不同，此外其值还依赖于 $(X_{ij}-\bar{X}_{\cdot j})$ 值。各学校内学生之间差异的所产生的随机影响 r_{ij} 也可以分离出来。具体而言，u_{0j} 是学校回归截距的随机误差，而 u_{1j} 是学校回归斜率的随机误差。对于这两个学校层次的随机误差，可以进行方差分析，其结果可以分别反映学校回归线的截距或斜率上的未解释方差是否还有继续解释的余地，即是否可能寻找其他新的学校变量对其学校截距或斜率上的残余变异做进一步的解释。

当我们将所有各项展开，方程便取得了标准形式：

$$Y_{ij}=\gamma_{00}+\gamma_{01}W_j+\gamma_{10}(X_{ij}-\bar{X}_{\cdot j})+\gamma_{11}W_j(X_{ij}-\bar{X}_{\cdot j})+u_{0j}+u_{1j}(X_{ij}-\bar{X}_{\cdot j})+r_{ij} \qquad [3]$$

上述方程的实际意义为：

某校某学生成绩 ＝ ①所有学校成绩的平均数＋②某学校特征的效应
＋③学生特征的一般效应＋④某学校的特征和学生特征的互动效应
＋⑤学校平均成绩的随机波动＋⑥某学校学生特征效应的随机波动＋⑦某学生的随机波动

下面我们仅从上述通用分层模型中所包含的一些子模型来显示其分析能力的强大。上述通用模型中的七项分量中有四项可以估计出效应参数（即截距和斜率）的，有三项则代表不同的随机误差效应。所谓子模型是指事先将通用模型中的某些参数或随机误差设为 0，其结果便产生某一种简化的子模型。这样一来实际上等于模型根本不考虑这些效应。

所有的子模型都会出现微观误差 r_{ij}，所以不再讨论它。模型可以分为两大类，区分的标准是看模型中是否含有层次互动效应，即对斜率{B}中的系统调整项 $\gamma_{11}W_j\ (X_{ij}-\bar{X}_{\cdot j})$ 和随机调整项 $u_{1j}\ (X_{ij}-\bar{X}_{\cdot j})$。

当不包含它们时，被称之为**随机截距模型**。比如当子模型中只有对应截距{A}中的项目类型时，意味着模型根本不考虑个体单位的变量的系统性影响，因此宏观单位之间只存在因变量 Y 平均水平估计上的差异。当加入对应斜率{B}中的 γ_{10} 存在于子模型中时，便意味着不同宏观单位中的个体特征 $(X_{ij}-\bar{X}_{\cdot j})$ 的构成将产生作用。但是由于 γ_{10} 代表的只是就个体案例的一般性斜率，对于每个宏观单位都普遍适用，因此宏观单位的结果虽然表现为依赖本单位个体特征（X 轴方向变化），形成具有一定斜率的回归线，然而各单位的回归线之间都是平行的，其结果所反映的仍然只是宏观单位之间水平差异。因此这些模型实际上假设学校之间影响虽然存在差异，但差异是固定的，不随个人特征变化而改变。

但是一旦模型中包括了对斜率的系统调整项和随机调整项，模型便可以显示学校之间更加复杂的差异。

下面先从比较简单的随机截距模型开始介绍：

（1）当既不考虑学校层次自变量的效应、也不考虑学生层次自变量的效应时，模型便简化为**随机效应的单因素方差分析**(one-way ANOVA with random effects)：

$$Y_{ij}=\gamma_{00}+u_{0j}+r_{ij} \qquad [4]$$

这一方程定义某一个学生的成绩都等于在总平均成绩的基础上加上所在学校施加的随机效应，以及本人的随机效应。对应这一方程的方差分析则意味着，学生成绩的总方差被分解为学校和学生两个层次的方差：

$$Var(Y_{ij})=Var(u_{0j})+Var(r_{ij}) \qquad [5]$$

分层模型的方差分析与常规方差分析十分类似，所不同的是这种方差分析是一种随机效应模型，因为学校的效应被定义为随机效应。

对学校随机效应进行方差分析的目的是检查不同学校的平均成绩之间是否存在差异。比如，对于 u_{0j} 的方差分析检验如果显著，则说明学校之间在平均成绩上存在显著差异，因而应该考虑寻找可用于解释这一差异的学校特征变量，在一般性的总平均成绩基础上加以调整，以更好地反映学校之间在成绩水平上的差异[12]。

（2）如果在上一模型的基础上再考虑学校的具体特征（比如表示学校经济实力的财务资金量）对成绩的解释作用，模型变成下列形式：

$$Y_{ij} = \gamma_{00} + \gamma_{01} W_j + u_{0j} + r_{ij} \qquad [6]$$

$$(\text{注：可视为} Y_{ij} = \beta_{0j} + r_{ij})$$

这种子模型称为**以(具体学校的成绩)平均数作为估计结果的回归模型**((regression model with means-as-outcomes)，因为这一回归的估计结果是学校截距，即 $\hat{Y}_{ij} = \beta_{0j}$，而学校截距在微观模型中又定义为学校平均成绩。

这一模型的主要任务：第一是检验模型中所定义的宏观自变量的作用 γ_{01} 是否显著，并描述这一作用的方向与强度；第二是在控制了模型中已定义的宏观自变量的条件下完成学校截距随机效应 u_{0j} 的方差分析检验。如果 u_{0j} 的方差不显著，意味着学校之间的成绩差异主要是由于学校财力差别造成的，此外已经没有明显的差别。如果 u_{0j} 的方差仍然是显著的，便说明学校之间还存在着财政财力这个自变量所不能代表的其他方面的差别（如教师水平、学校风气等）。

这个模型经常被用于检查组织、单位、地区等宏观特征影响的研究。注意这一模型不同于完全忽略个人变异效应的简单的"生态学"模型，因为它在本质上是多层模型，并没有忽略反映个人随机变异的 r_{ij}。此外，注意这一模型所反映的变量是集合层次的特征，因而又不同于下面将要介绍的协方差分析模型，因为后者所控制的自变量是个体层次的特征。并且，这一汇总层次的变量既可以是分类性质的（比如将 W_j 定义为表示是否为重点校的虚拟变量），也可以是连续变量，如本例中用的是学校财务资金量。在模型中明确地定义出学校变量 W_j，模型便可以估计出这个变量相应的回归系数 γ_{01} 来。

（3）如果想检查在控制学生特征（比如学生的智商这样的连续变量）的条件下不同学校是否对学生成绩还有显著作用，分层模型将成为：

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}..) + u_{0j} + r_{ij} \qquad [7]$$

虽然这一模型中并没有明确定义学校变量，但仍然可以通过表示不同学校的下标来进行学校的分类，比如当我们并不在意学生所处的具体学校，而只关注其所属学校是不是重点学校（j=0 表示普通校，j=1 表示重点校），那么分层模型便成为**随机效应的单因素协方差分析**(one-way ANCOVA with random effects)（图 1）。注意现在学生变量 X 是以总平均值为中心测量的，而所在学校分为两类。于是，这一子模型代表了两类学校各自的回归线。协方差分析的目的实际上是在控制学生变量影响以后再来检查学校类型之间是否还有成绩上的差别。也可以说，它适用于检验"简化论理论"是否与实际相符合。比如，当我们希望澄清重点学校之所以成绩好是否仅仅是因为其学生素质更好的问题时，便用到这种模型。如果在控制学生智商的条件下学校随机效应 u_{0j} 的方差不显著，便证实了上述怀疑。否则，意味着两类学校在学生素质差异之外，还有其他学校方面的原因促使学生取得较好成绩。此外，该模型也同时测量并检验学生智商对成绩的影响。注意，虽然代表不同类型学校的两条回归线的截距水平有差异，但是回归线是平行的，因此它还属于随机截距模型，其模型的前提假定是，在两类学校中取任何相同智商学生来对比时，其差别不随智商的变化而变化。这是模型设置所决定的性质，而并不一定是真实情况的反映。

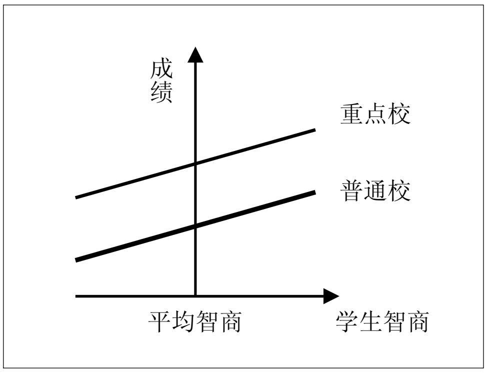

图 1　分层协方差分析的回归线

（4）分层模型在进行协方差分析时还可以同时检验宏观协变量和微观协变量的作用，以适应更一般的

[12] 由于篇幅所限，本文没有涉及对个人随机效应 r_{ij} 的进一步分解及其方差分析，而这在分层模型中也可以进行。

情况。在假设没有层际互动效应的情况下，这种扩展的随机效应协方差分析的组合模型为：

$$Y_{ij}=\gamma_{00}+\gamma_{01}W_j+\gamma_{10}(X_{ij}-\bar{X}..)+u_{0j}+r_{ij} \quad [8]$$

上述方程在形式上类似于一个多元回归，但是常规的回归分析中不能同时检查不同层次的自变量，而这一模型却可以根据分层结构数据同时求出各种效应的参数估计。尽管这个方程已经是多元方程，但图 2 仍用二维坐标图来表示，而学校财力这个连续变量则采用离散的系列回归线表示。财力大学校的回归线处于最高位置，表明学校的高投入有助于学生成绩的提高。虽然图 2 中不可能为所有不同财力水平的学校分别画出具体的回归线，但我们知道它们之间的成绩差异幅度是与学校投入成正比的，并且所有的回归线之间都是平行的。

对这一模型的方差分析是在将宏观变量与微观变量一起作为协变量的条件下进行的，而常规协方差分析则做不到这一点。不仅如此，实际上分层模型还可以摆脱常规方差分析、协方差分析和回归分析都必须遵守的同质性假定(homogeneity，即假设在不同类型中或在不同自变量值情况下有因变量方差相等)。

此外，在这里还需要强调：分层模型在不同层次采用同源的自变量也是可以的。比如，学生层次的自变量为个人智商，而学校层次仍然可以采用由学生层次汇总的平均智商（或学生智商的方差等其他同源的汇总变量）作为自变量。这时，模型对学校自变量作用的测量和检验，实际上已经是在检验“昔孟母，择邻处”背后依据的理论是否真的反映实际情况。这与测量和检验学生个人层次智商对成绩的影响有完全不同的意义。

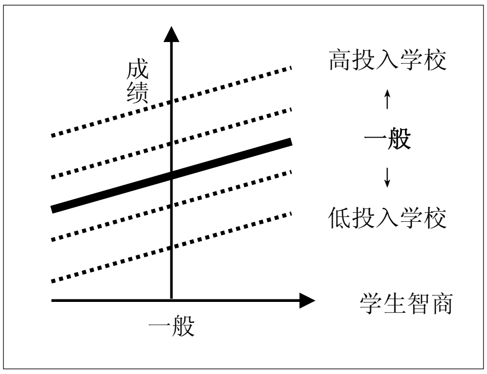

图 2　同时考虑不同层次自变量的模型
（假设无层际互动效应）

以上几种子模型都属于随机截距模型，而以下介绍的两种模型则强调了自变量对成绩的作用斜率在各学校之间是可变的，这样一来，便可以检查个体自变量的作用模式是否依所属不同群体而发生变化。实际上，探究不同群体中的不同模式正是分层模型的主要关注点。

（5）随机系数回归模型(ramdom-coefficients regression model)的主要功能同时检查学校之间的回归系数（包括截距和斜率二者）的模式是否存在差异。在假设学校层次的两个基本模型（式 2a 和 2b）中的学校自变量W_j的系数都为 0 的先决条件下，得到的组合方程为：

$$Y_{ij}=\gamma_{00}+\gamma_{10}(X_{ij}-\bar{X}._j)+u_{0j}+u_{1j}(X_{ij}-\bar{X}._j)+r_{ij} \quad [9]$$

在这个模型中作为学生特征$(X_{ij}-\bar{X}._j)$的斜率系数的效应有两项：一项是其适用于所有学校的固定效应γ_{10}，那么模型主要是求出该参数估计，并检验其是否显著。另一项则是该模型的特色所在，它就是对应各具体学校回归线斜率的随机效应u_{1j}，由于它的存在该模型便不再属于随机截距模型的范围了。对这一效应进行方差检验可检查不同学校的斜率是否存在显著变化。这一学校回归线斜率的随机效应，在前面几个子模型中尚未出现过，原因是那些模型都假设学生特征作用的斜率在各学校中不变。注意以上随机截距子模型的公式中，或者根本不考虑学生变量，或者对学生变量的测量是以总均值为中心的；而随机系数回归模型对学生变量的测量却是以学校的均值为中心的。

对于u_{1j}的方差分析如果显著，便肯定了学生特征对成绩的作用斜率在不同学校之间的确存在着统计性显著的差异，尽管模型并未对此差异做出具体估计。与此同时，对u_{0j}的方差分析还可以检查在已经考虑学生个人智商的因素以后各学校平均成绩（即回归线截距）上是否仍存在差异，以了解是否还需要寻找学校之间在其他方面的学生差异来解释学校平均成绩上的差异。

（6）将截距和斜率作为估计结果的回归模型(regression model with intercepts and slopes-as-outcomes)则可以将学校之间在学生个体特征与成绩回归线的截距和斜率上的差别形式的假设原因在模型中明确定义出来，以便求出有关参数估计并进行假设检验。这种模型便是方程 3 所示的完整的通用模型：

$$Y_{ij}=\gamma_{00}+\gamma_{01}W_j+\gamma_{10}(X_{ij}-\bar{X}._j)+\gamma_{11}W_j(X_{ij}-\bar{X}._j)+u_{0j}+u_{1j}(X_{ij}-\bar{X}._j)+r_{ij} \quad [10]$$

(注：该模型的称呼源于上式可视为 $\hat{Y}_{ij}=\{\beta_{0j}\}+\{\beta_{1j}\}(X_{ij}-\bar{X}._j)$)

与上一随机系数回归模型相比，完整模型又多了两项系统效应，即$\gamma_{01}W_j$和$\gamma_{11}W_j(X_{ij}-\bar{X}_{\cdot j})$。其中，$\gamma_{01}$反映了学校变量$W_j$对一般性回归截距$\gamma_{00}$的改变，$\gamma_{11}$则反映了两个层次上自变量之间的互动$W_j(X_{ij}-\bar{X}_{\cdot j})$对一般性回归斜率$\gamma_{10}$的改变。如果这两项学校层次对回归系数的调整效应的参数估计统计检验显著，便意味着不同学校的回归线的确存在着截距和斜率上的差别。

如图 3 中的(a)图所示，不同学校回归线在一般智商情况下的“截距”（表示中等智商时的回归线高度）有所差异。并且，在一般情况下（由粗线表示的“一般”回归线），学生智商高会有较好成绩。但是不同学校中学生智商与成绩的关系存在模式上的差异。比如学校 1 的模式是好上加好，即智商较差的也能取得好成绩，而智商较高的成绩更好。而学校 2 则反映出马太效应，智商低的成绩很差，智商高的成绩很好，这种学校显然更适合于聪明学生，而牺牲了接受能力弱的学生。学校 3 则显得教学主要“偏向”于那些接受能力较慢的学生，促使他们的成绩高于一般水平，但这是以牺牲其他学生为代价的。学校 4 则表现为教学效果很差，但并没有出现明显的重视一头的“偏向”。了解学校之间的这些差别对于不同智商的学生对不同学校的选择是很有意义的，对于进一步理解学校的行为模式及其背后更深层的原因也是有意义的。

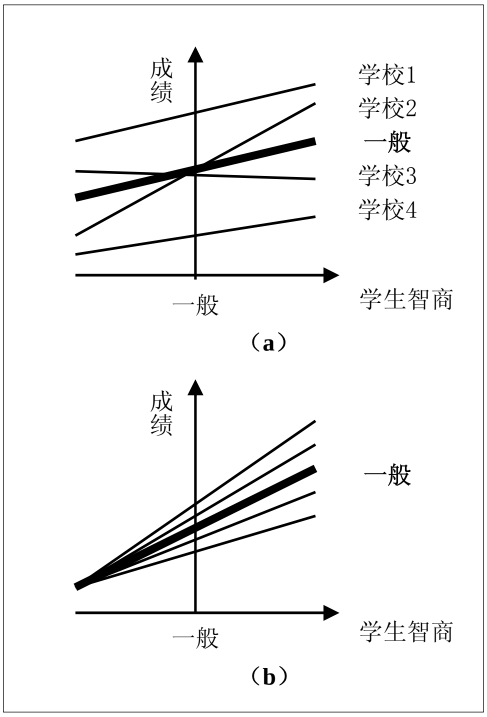

图 3　随机系数回归模型

上述图示中的不同学校行为模式实际上揭示出一个具有普遍意义的重要命题，公平与效率。Raudenbush 和 Bryk（2002: p18）在其分析中，便明确地提出可以将分层模型的回归截距看作效率的标志，而将分层模型的斜率作为公平的标志[13]。

有时，研究可能关注，相同个人特征的学生在学校之间的差异分布会不会随个人特征取值的不同而有所变化。图 3(b)描述的便是这样一种情况：各学校都没有从本质上改变学生智商越高成绩越好这样一种一般关系。但是，可以看出对于不太聪明的学生而言，学校之间的差异较小；而对于很聪明的学生而言，学校之间的差异很大。这种差异说明选择学校的重要性对于不同特征的个人是有差别的。Duncan 等人也用健康研究的例子说明（1998：99-100），发病率为年龄的函数是一般规律，但是对于某一年龄（如老年人）居住于不同社区可能会导致很不同的发病结果，而不同社区的年轻人则没有什么明显差别。

在将截距和斜率作为结果的回归模型中还可以对影响学校回归截距和斜率的两项随机效应u_{0j}和u_{1j}进行方差分析检验。这时的检验是为了分别检查在学校回归截距和斜率的残差方差中是否还有进一步解释的余地。当方差分析结果不显著时，说明在残差方差已没有进一步解释余地了，如果显著则意味着还可以寻找其他学校层次变量来进一步对学校之间的差别进行解释。这意味着，学校层次的截距或斜率模型还需要向多元自变量方向扩展。

以上对分层模型的几种子模型的讨论主要是为了介绍分层模型的主要功能。为了更加形象化，采用了一些图示。实际上，分层模型在探索不同行为模式时，并不需要直接将所有单位的回归线都画出来看，因为要是有很多单位时，这种图便很难看了。因而它采用更有效率的统计方式，即对各单位回归线的两个参数（截距和斜率）之间计算相关的方法或画出相关散点图（一个单位在此图上只是一个点）的方式来检查，当相关程度高时则表明单位之间有比较一致的行为模型，否则意味着行为模式在单位之间有明显不同。这与我们通常用散点图来表示 X－Y 关系是一样的，只不过在这里看的是学校层次回归线的截距和斜率之间的关系。在分层研究中，学校也是某一层次中被抽出来的案例。

在上述介绍中，我们可以总结出，对于 u 类学校随机效应的方差分析可以反映学校层次未解释的方差是否足够大，它在实际研究中的主要意义是了解是否有必要在下一步研究中继续寻找具体的学校层次的自变量来“解释”或“确定”本模型所尚未解释的学校方差。因此，它可以视为在具体定义和估计新的学校

[13] 由于他们所用的例子是学生社会经济背景对学习成绩的影响，因此认为回归斜率等于0便是公平，表示“有教无类”。而本文所涉及的自变量是个人智商，那么显然同样的教育投入带来的成绩产出是不一样的，所以恐怕用“一般”斜率作为公平的标志更恰当一些。

层次自变量之前的预备研究。实际上，对于学生层次的随机误差也同样可以做进一步的分解和分析。

总之，分层模型通过将个体因变量变异分解到不同层次的水平效应（截距）、变量效应（斜率）、和不同随机误差上去，其通用组合模型中的各项效应代表了这种模型可以研究分析的不同方面。在实际应用中，可以根据不同研究假设的需要，省略与研究不相干的那些项目，采用简化了的子模型。

这一模型最大的贡献是可以估计和比较不同层次效应的解释作用，比如回答诸如“到底是学生因素还是学校因素对学习成绩的作用更大”这类的问题。因此，过去只能对单一层次分析的限制已经被破除了，于是过去由于这种限制而产生的只强调个人因素及其构成特征的简化论观点和只强调环境和组织因素的生态学观点的争执，现在便有可能根据这一新模型来定量化地在多层框架中加以整合的具体分析。

该模型分析方法的特点是，它不仅可以提供各类效应的参数估计，实际上还特别注重个体在特定处境中行为过程的不同形式。这种分析在常规分析中由于同质性假定受到很大限制，然而这种分析在分层模型中则被视为例常的情况来对待。

分层模型另外一个优点是包容性很强，以上所介绍的若干主要子模型，只体现了其包容性的一个方面。另外，这种包容性还表现在每一层的自变量数可以扩展为多元，并且层次的数目也可以扩展为多层。

近 10 年以来,分层模型又取得了长足的发展。其重要标志是这一方法的代表著作《分层线性模型：应用与数据分析方法》(Raudenbush 和Ｂryk, 2002)已经出版了第二版[14]。该书的第二版增加全新的 4 章（200 多页）内容，主要反映了分层模型的包容性的进一步扩展。分层模型的新扩展可以容纳许多不同类型的分析技术，比如在分层模型中既可以用线性回归、也可以用 logistic 回归来表达某一层中因变量和自变量的关系。换句话说，就是各层模型的因变量既可以是定距变量、也可以是二分类的、多分类的定类变量或者是定序变量。并且，这种扩展还包括了现在的分层模型已经可以处理测量误差和潜在变量。此外，这种模型所能应用的多层数据结构的形式也十分丰富，包括在同一个人之下可以嵌套其在不同时点的追踪观测值（这意味着时间变化也可以被容纳到分层模型中来），也包括同一个人还可以双重地嵌套于不同的上级层次（比如居住社区和工作单位）。另外一个方面的扩展是分层模型可以应用贝叶斯推断估计，这一发展使分层模型方法摆脱了以往需要各层次都有较大样本规模的要求，现在也可以处理相对较小样本规模的问题，并且便利了对某一具体分组单位的估计，这意味着分层模型分析进入了具体单位绩效评价的阶段[15]。更重要的是，分层模型近年来的发展不是仅涉及到概念和原理方面，还解决了计算程序问题，意味着这些扩展已经达到了普及应用阶段。

参考文献

艾尔·巴比．2000．《社会研究方法》（邱泽奇译），华夏出版社．

弗庭格尔． 1995．家庭生命周期的统计测量．载约翰·邦加茨等主编，《家庭人口学：模型及应用》，北京大学出版社．

郭志刚．1999a．北京市家庭户规模的分解研究，《人口研究》第 3 期．

郭志刚．1999b．北京市早婚情况分析，《中国人口科学》第 3 期．

郭志刚、邓国胜．1998．年龄结构波动对婚姻市场的影响，《中国人口科学》第 2 期．

王治平．2002．《经济学阐释》，中国物价出版社．

Raudenbush, S. W. and A. S. Bryk. 2002. ***Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition***. Advanced Quantitative Techniques in the Social Sciences Series. Sage Publications, Inc.

Duncan, Craig, Kelvyn Jones and G.raham Moon. 1998. Context, composition and heterogeneity: using multilevel models in health research. ***Social Science and Medicine***, Vol. 46, No.1.

[14] 该书第一版出版时作者署名顺序是Bryk和Raudenbush(1992)。

[15] 过去这种分析主要是为了把握各分组单位的总体的分布特征。