期末报告 - 多层线性模型(MLM)方法操作介绍

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus.* Routledge (chapter1-chapter3)

20162231 林凌真 2018/1/20

1 引言

在心理学研究中，广泛存在着数据结构带有层级嵌套特性的情况。这一层级结构性一方面可能直接表现在理论构建中，感兴趣的研究变量本身就属于不同的理论层级，例如工业组织心理学中有些变量来源于个体水平、有些变量则属于部门、公司水平；另一方面，也可能出现在数据层面，许多时候由于并不完全随机的取样方法，比如在某学校、某社区附近收集数据，则存在着学校中有班级、班级内有个体，社区内有楼、楼内有住户这样的嵌套关系。

在面对层级结构数据时，分析方法也必须相应的将数据由于层级结构而存在的内部不独立性考虑进来。在不同的研究领域下，这样的方法具有不同的术语，比如随机系数模型、混合效应模型、多层回归模型、层级线性回归、多层结构方程等，本文将这些相似的概念皆统称为广义的多层模型。这些方法的核心是相同的，即将单一结果变量或是多个结果变量的方差拆解，通过一系列不同层级的解释变量来解释所拆解的结果变量的方差。

值得注意的是，量化研究的用途往往被认为有预测和解释两种不同的方向。在预测性的研究中，变量最终是否存在基本取决于统计上是否显著，为了预测效果的最大化，不显著的变量将被直接从模型中去除。而在解释性的研究中则不同，研究者首先构建理论模型来解释某些现象、结果，接着猜在数据中研究这一模型的表现是否符合实际。多层模型分析一定是解释性大于预测性的，这是因为这个分析方法的本质和内在逻辑——若是为了预测效果最大化，完全不考虑所谓层级直接考虑每个预测变量对结果变量的解释率是最稳定可靠的，层级本身就是出于理论的角度才加入分析。因此研究者在进行多层模型缝隙时，一定要不断强化理论的重要性，只考虑模型拟合与参数估计显著与否很可能导致研究结果不伦不类，既无法解释亦无法预测。

本文将首先介绍多层线性模型(MLM)的分析方法，具体拆分为四个分析步骤。接着用 Mplus 作为分析软件，结合实例一步一步介绍分析时的代码。

2 多层线性模型(MLM)的分析方法

多层模型的原理在于，将单层模型中预测变量的系数由固定变为随机。简单来说，固定系数意味着预测变量与结果变量之间的关系在所有个体当中都是相同的。而由于层级结构数据中，属于同一个高一层结构的小群体之内存在相关，相对应的就会与其他小群体之间存在差异（比如同一个班的学生积极性相近，而不同班的学生之间差异较大），因此系数不再设定为所有个体相同，而是随着层 2 水平的变化而变化。下面首先以最基础的一元线性回归为例来看看，系数随机化这一从单层分析到多层分析变化的过程，并简单总结单层模型和多层模型的区别。然后介绍具体多层模型建立的四个步骤。

2.1 从单层模型分析到多层模型分析

让我们假定，现在的研究主题为探究社会经济地位（SES）对学生数学成绩的影响。以 SES 为横坐标、数学成绩为纵坐标作散点图如图 1 所示。

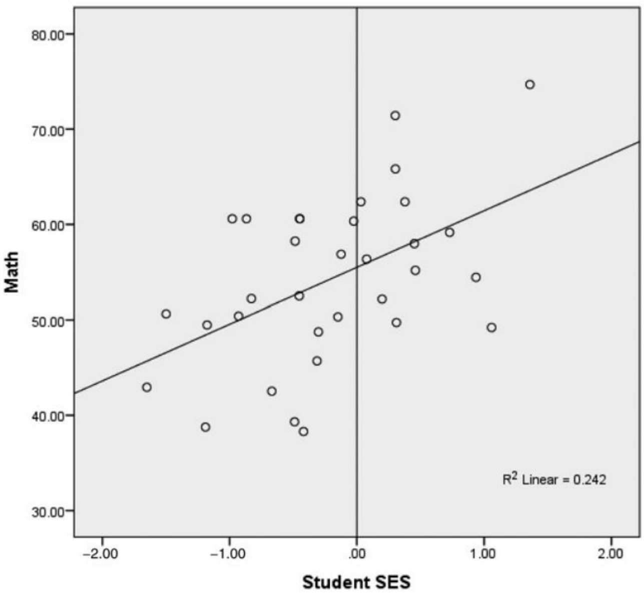


图 1 散点图 - 社会经济地位 SES 与学生数学成绩的关系 & 拟合得到的回归方程

根据最小二乘法，构建一条使得真实数据值与预测值之间差值的平方和最小的直线，即为一条有截距和斜率的非标准化回归方程：$Math_i = \beta_0 + \beta_1 SES_i + \varepsilon_i$。对第 i 名学生，它的数学成绩$Math_i$将等于截距$\beta_0$加上斜率$\beta_1$乘以他的社会经济地位$SES_i$加上测量误差 ε_i。社会经济地位对数学成绩的解释能力越强，所有个体的测量误差方差总和将越小，极端情况下所有观测点都将落在回归方程对应的线上，也即误差方差为 0。在本例中，得到的方程为：$\widehat{Y_i} = 55.49 + 5.94\, X_i$，截距为 55.49，斜率为 5.94。可以看出，此时截距和斜率都是固定值。事实上，所有的单层回归方程，所得参数都将是固定值。而误差项虽然不体现在这个方程中，但它

的形式也是固定的：所有人的误差项将形成一个正态分布，独立于 x 不随 x 的变化而变化，均值为 0，方差同样是一个固定值。

那我们回到这个例子本身。如果我们回收的数据，其实来自于 5 所不同的学校，那么同一个学校内的学生很可能 X 与 Y 的关系会与另一所学校的学生不同。也就是说，如果给 5 所学校分别拟合出一个回归方程，最终得到的 5 个截距、5 个斜率很可能彼此之间差别很大，如图 2 所示。此时，我们所做的事就不再是参数固定值了，而是允许参数随着学校的不同而不同，每个学校能得到属于自己的独特参数，这就是参数随机化。

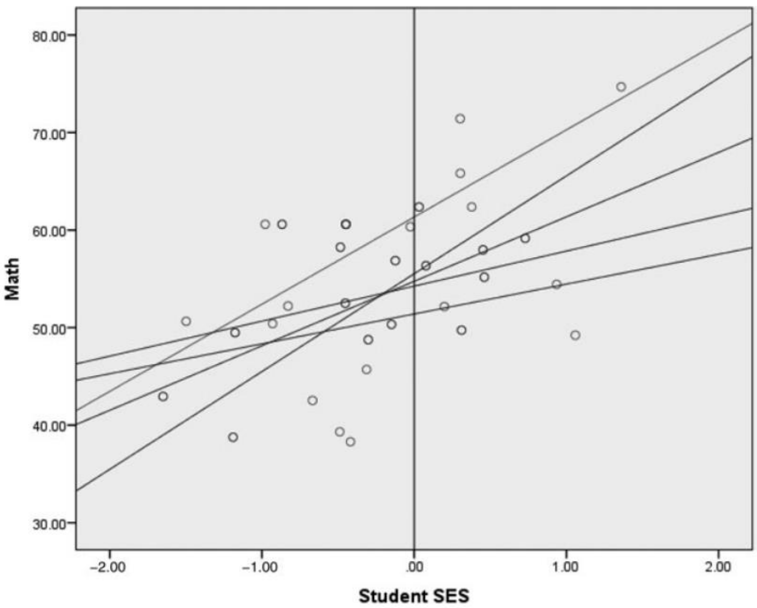


图 2 散点图 - 社会经济地位 SES 与学生数学成绩的关系 & 5 所学校分别拟合 5 条不同的方程

值得一提的是，如果这 5 所学校所得的 5 条线并没有区别，而是基本重合在一起，那就证明学校这个层级的存在并不会对学生水平测得的数据产生影响，那就说明参数随机化是没有必要的，此时就不要再使用多层模型的分析来增大复杂程度了，直接使用单层分析即可。当然，这是两个参数截距和斜率都没有显著差异的情况，实际应用中更多的情况是斜率跨学校没有差异而截距有差异，此时所采用的就是随机截距的多层线性模型，由于是最常见的情况往往省略随机截距这个定语。相对少见的情况是截距与斜率都随着学校的不同而不同，此时则会特别标记为随机斜率的多层线性模型。

总结来说，多层模型相比单层模型，(在合适的情况下适用)有以下几处优势以及可行性：

- 提供更符合理论层级的分析，尤其是在探究不同层级变量之间关系的情况下，使用多层模型能把变量置于本来的水平上分析

- 更恰当的处理误差。当数据内部存在一定比例的层级结构时，允许参数随着层 2 水平的变化而变化能使得误差减小，参数估计更准确
- 正确的估计标准误。使用单层模型估计带有层级结构的数据，会导致低估标准误，进而可能得到假性显著的结果，增大一类错误的概率
- 多层模型现在的灵活性较高，能够拟合一系列不同的模型
- 经过多年的研究，如果多层模型的拓展丰富，各类数据结构假设模型基本都可以验证
- 各种软件的开发使得如今分析层级数据十分便利

2.2 多层模型建立的四个步骤

在本例中，计划研究的问题是哪些组织层面和个体层面的因素会影响公司内部成员的“生产力”。假设个体生产力分别会受到个体层面的因素“动机”以及部门层面的因素“资源”的影响，另外假设这个多层模型中截距&斜率都是随机的。假设模型如图 3 所示。

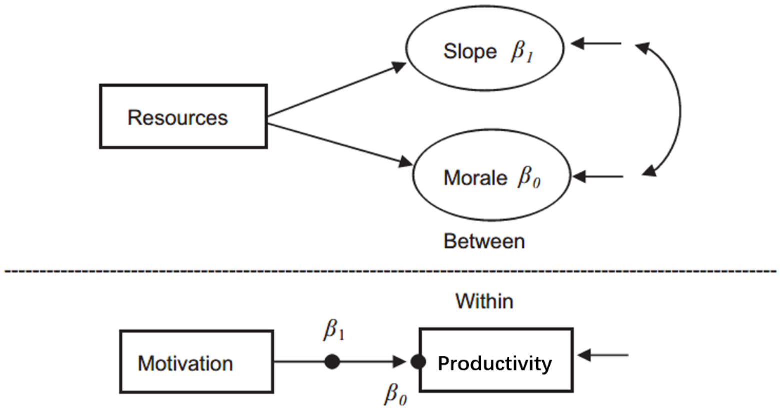


图 3 假设模型图 - 随机斜率的多层线性模型（个人动机&组织资源水平影响个体生产力）

模型建立分为以下 4 步： 1.将结果变量的方差划分为组内与组间（此时已经是随机截距模型，只是没有预测变量）； 2. 添加水平 1 的预测变量来解释结果变量的组内方差； 3. 确定该模型的斜率是否需要随机； 4. 添加水平 2 的预测变量来解释结果变量的组间方差以及随机斜率的方差。

2.2.1 步骤一：将结果变量的方差划分为组内与组间

是否有必要进行多层模型的分析，一定程度上取决于因变量是否具有层级结构。只有当因变量本身就既有组内差异又有组间差异，才能够通过使用水平 1 预测变量和水平 2 预测变量分别解释这两个不同的成分。在这一步骤中，只考察结果变量，是有层级结构的假设，但不加入任何预测变量的。此时的回归方程为：

$$\text{Level-1} \qquad Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

$$\text{Level-2} \qquad \beta_{0j} = \gamma_{00} + \mu_{0j}$$

在水平 1 中，以员工为单位，每个员工的生产力水平都围绕所属部门的生产力平均水平上下浮动。在水平 2 中，则是部门的生产力水平相对应的围绕总体生产力均值而上下浮动。具体来说，第 j 个部门的第 i 个员工，他的生产力水平Y_{ij}等于他所属的部门平均生产力β_{0j}加上个体水平的误差项ε_{ij}。而这个部门的生产力β_{0j}又会等于总体平均生产力γ_{00}加上部门水平的误差项μ_{0j}。将这两个方程合并后得到$Y_{ij} = \gamma_{00} + \mu_{0j} + \varepsilon_{ij}$，方程中的$\beta_{0j}$没了，剩余 3 项预测量分别是总体均值(即截距项) γ_{00}、组间误差μ_{0j}、组内误差ε_{ij}，其中第一项截距是固定值，而后两项是根据实际数据的变化而变化的可知参数，对其分别取平方和分别可得组内误差和组间误差的大小。因此则有：

$$Var\left(Y_{ij}\right) \;=\; \mathrm{Var}\left(\mu_{ij} + \varepsilon_{ij}\right) \;=\; \sigma_b^2 + \sigma_w^2$$

结果变量的方差（变异量）就通过这个公式，被拆分成了组内变异和组间变异。此时就可以计算组内相关(intraclass correlation, ICC)，即组间方差占总方差的比例，公式如下：

$$\rho = \;\; \sigma_b^2 \,/\, (\sigma_b^2 + \sigma_w^2)$$

此时，我们可以决定是否应该使用多层线性模型来替代单层线性模型分析数据了。这一决定应该同时考虑理论层面和数据层面，具体分为四点：1. 结果变量是否在理论模型中本身就具有层级结构；2. 水平 2 的样本量是否足够大；3. 协同考虑 ICC 与水平 2 平均的组的大小（比如平均每个部门内有多少人）；4. 过往研究中是否采用了多层的方法。综合考虑这四点后，认为有必要应用多层方法，则必须使用否则会导致统计上一类错误的概率增大或是不匹配理论模型，这二者都是严重的研究问题。

总的来说，这一步骤是将处于水平 1 的结果变量的方差拆分为两个部分，也即进行了因变量的随机截距的过程，这里讲这一过程可视化为图 4 中红框标识的内容。图的下半部分表示组内的变量情况，上半部分为组间的变量情况。此时模型中仅有一个变量，就是结果变量生产力，虽然图中有两个框/圈，但实际上只是表示同一个变量的两个成分。另外，在组内部分因变量外框上的一个红点表示的是截距β_0的随机化。

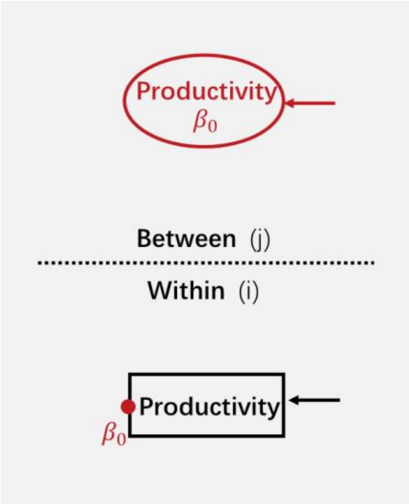


图 4 步骤一：将结果变量的方差划分为组内与组间(随机截距化)

2.2.2 步骤二：添加水平 1 的预测变量来解释结果变量的组内方差

在这一步骤中进行的分析如图 5 所示，简单来说就是增加了一个层 1 的预测变量“动机”。

$$\text{Level-1} \qquad Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$

$$\text{Level-2} \qquad \beta_{0j} = \gamma_{00} + \mu_{0j}$$

此时，步骤一中的回归方程的水平 2 不变，依然是层 1 的截距围绕层 2 的部门变化，但层 1 的方程中要加入第一个除Y_{ij}以外的变量X_{ij}，并把这一项的斜率记为β_{1j}。由于在此步骤中，设定斜率恒定，因此也记做没有下标 j 的β_1。此时结果变量的组内误差方差中的一部分被预测变量所解释，而组间方差则不受影响。

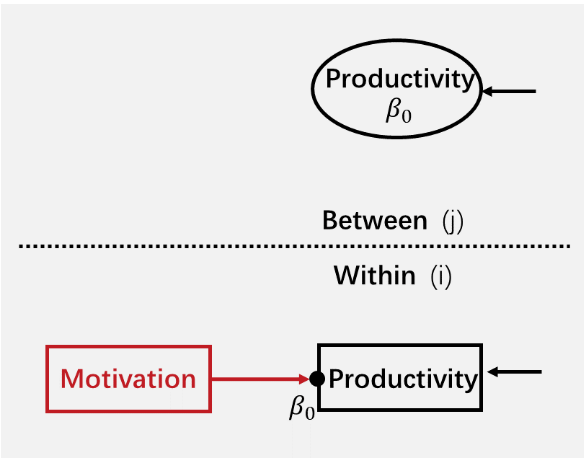


图 5 步骤二：添加水平 1 的预测变量来解释结果变量的组内方差

正如前一部分中所提到的，多层模型在处理数据时就是不再把整个数据集中的所有观测值一起分析，而是将总样本分为 j 个子集，在每个子集中进行预测变量和结果变量关系的参数估计，因此可以简单理解为：可以得到 j 个截距β_{0j}、j 个斜率β_{1j}。若理论认为“不同 level-2 unit 的值不同” or “这 j 个截距β_{0j}之间差异很大”则应使用随机截距模型，这是在步骤一当中所进行的。此外，若理论认为“X→Y 的效应随 unit 的变化而变化” or “这 j 个斜率β_{1j}之间差异很大”则应使用随机斜率模型，这将会是步骤三的内容。总结来说，多层模型的核心就是参数随机化，而是否进行参数随机化、对哪个参数进行随机化取决于：1. 理论基础；2.层级二的方差。

此外，值得一提的是参数的实际意义。在方程中，截距项的意义一般为当预测变量为零时结果变量的水平，本例中估计得到的γ_{00}表示当员工的动机为 0 时，生产力的水平。斜率的意义为预测变量每变化一个单位，结果变量随之变化的数值、在本例中估计得到的β_1表示员工的动机每增大 1 个单位，生产力水平增大的幅度。为了截距项的解释，在实际应用中就要使预测变量为零有实际意义，通常的做法是对预测变量进行中心化，此时截距的意义就变成了“当员工的动机为平均水平时，他生产力的水平”。

2.2.3 步骤三：确定该模型的斜率是否需要随机

如同前文所述，关于这个模型当众的斜率是否需要随机并不是固定的，而是要取决于实际数据或是研究前的理论依据。值得一提的是，由于设定模型中的斜率随机是一个很复杂的操作，尤其是当水平 1 中有多个预测变量指向结果变量（因此有多个截距）时，设定斜率随机会对模型的估计产生较大的压力，因此一般只在理论需要并且认为这一随机性极大的情况下，才会预先设定模型中存在随机斜率效应。

$$\text{Level-1} \quad Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$

$$\text{Level-2} \quad \beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$

从回归方程可以看出，与步骤二相比只是多了一个层 2 的方程，将层 1 的斜率β_{1j}也像截距一样设置成了带有误差项的围绕总体均值波动。从图 6 则可以看出，这个步骤就是简单的添加了一个成分：水平 2 的斜率，而这个成分的产生就是让层 1 的斜率随着层 2 而变化了。而这个斜率因子与步骤一、二中产生的截距因子之间可以存在相关（也可以没有）。

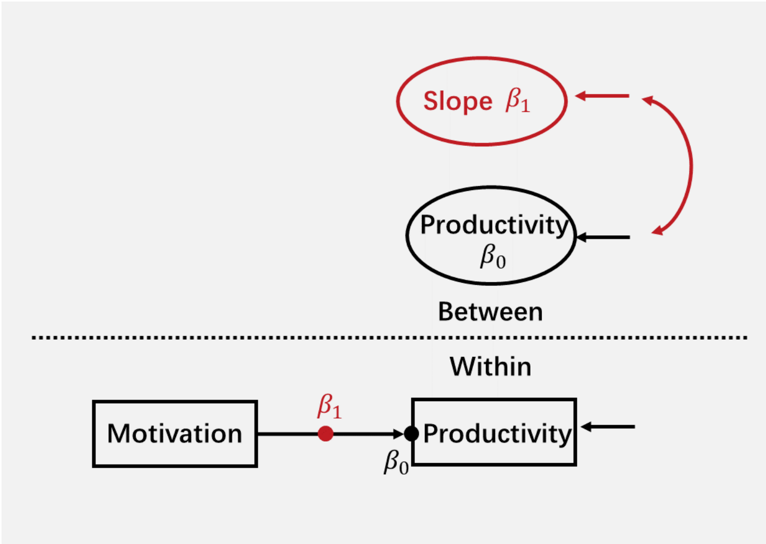


图 6 步骤三：确定该模型的斜率是否需要随机

2.2.4 步骤四：添加水平 2 的预测变量来解释结果变量的组间方差以及随机斜率的方差

在理解了前三个步骤的前提下，第四个步骤就显得格外简单——加入水平 2 的另一个或几个预测变量，来解释步骤一中划分出来的结果变量的组间方差，另外也加入变量解释步骤三中设为随机的层 1 方程中的斜率。在本例中，解释此二者的都为同一个部门水平的变量“资源”，记为W_j。与步骤三相比，回归方程的变化很小，仅仅是在层 2 的两个方程中各加入了一个预测项。在图 7 中依然用红线条标记了这个步骤所做的事。

Level-1 $\quad Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$

Level-2 $\quad \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j}$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + \mu_{1j}$$

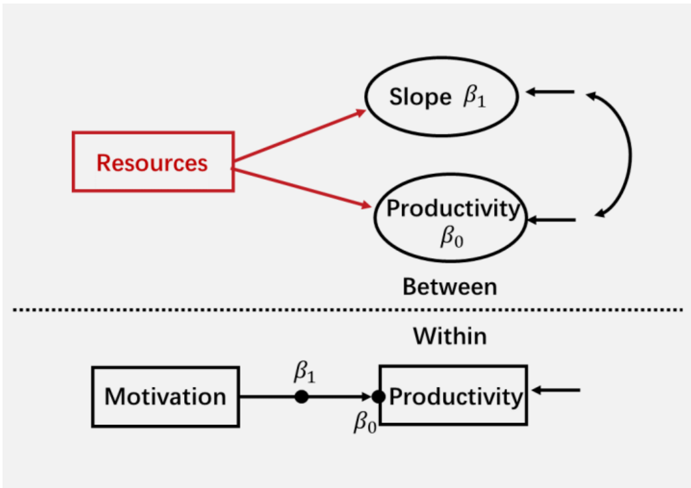


图 7 步骤四：添加水平 2 的预测变量来解释结果变量的组间方差以及随机斜率的方差

至此，一个最基本但完整的多层线性模型的建立步骤已经呈现完毕，这个示例中包含了两个不同水平的预测变量 X 和 W，以及一个处于水平 1 的结果变量 Y，且 Y 的水平会随着层 2 单位的变化而不同（即截距随机）。另外，变量 X 对变量 Y 的作用同样会随着层 2 单位的变化而变化（即斜率随机）。特别要注意的是，这个 X→Y 的效应之所以变化，方程中水平 2 的变量 W 能够解释，也就是说水平 1 内部的组内效应 X→Y 被水平 2 的变量 W 所调节了。这个跨层级的交互作用由回归方程来看可能更加的直观。

让我们将得到的 3 个最终的方程代入为一个，则会得到：

$$Y_{ij}=\gamma_{00}+\gamma_{01}W_j+\mu_{0j}+(\gamma_{10}+\gamma_{11}W_j+\mu_{1j})X_{ij}+\varepsilon_{ij}$$

将这个方程各项乘开，并调整顺序后会得到：

$$Y_{ij}=\gamma_{00}+\gamma_{10}X_{ij}+\gamma_{01}W_j+\gamma_{11}W_jX_{ij}+\mu_{1j}X_{ij}+\mu_{0j}+\varepsilon_{ij}$$

这个方程中一个有七项，代表着 7 个需要估计的参数，它们可以根据是否为固定值划分为两类。其中前四项分别为：截距项γ_{00}、组内效应$\gamma_{10}X_{ij}$、组间效应$\gamma_{01}W_j$、跨层级交互作用$\gamma_{11}W_jX_{ij}$，截距项和后三项的系数（斜率）都是可以估计得到的固定值。而后三项则为随机效应，分别代表了随机截距与斜率的共变/相关程度、组间误差和组内误差，通过估计方法可以得到他们的协方差或是方差。通过观察这一方差可以知道，在截距与斜率都随机的多层模型中，包含了一个组间效应、一个组内效应以及一个跨组的交互效应。而只有随机截距没有随机斜率的方程，则没有交互效应。至此，完整的分析已经介绍完毕。

3 MLM 实例数据分析&相应 MPLUS 程序代码介绍

在这一部分中，将介绍另一个例子以巩固前一部分的模型构建步骤的理论基础。在这个例子中，想要探究的问题是：

1 员工对于薪水的满意程度将如何影响他们的工作斗志？（控制性别、种族）

2 所处部门的整体工资水平是否能在问题 1 中所述满意度&斗志的关系间起调节作用？

根据这两个假设以及上一部分介绍所提到的模型图画法，我们可以画出本例所对应的假设模型，如图 8 所示：

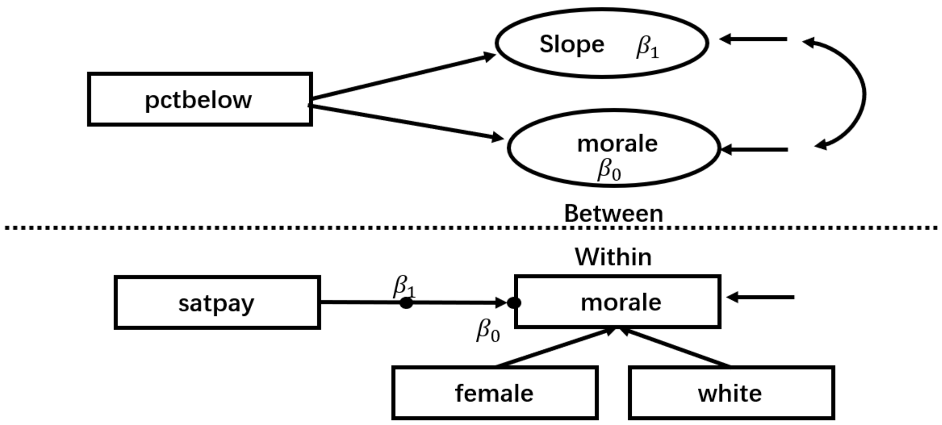


图 8 实例对应的假设模型图

收到的数据中共有 13189 名企业员工，分别来自于 165 个不同的部门。每个员工都报告了自己的薪水满意度(satpay)、工作斗志(morale)、性别以及种族(编码为 female、white 两个哑变量)，即水平 1 共测量了四个变量，其中包括一个预测变量、一个结果变量和两个控制变量。除此之外还收集了每个部门的工资水平相对于总体平均工资高或低的百分比(pctbelow)，即水平 2 仅有一个预测变量。描述性统计的信息见图 9。

	N	*Minimum*	*Maximum*	*Mean*	*Std. Deviation*
Departments					
Pctbelow	165	2.40	82.90	27.091	15.514
Individuals					
satpay	13189	0	16	8.882	3.415
morale	13189	11	40	26.306	6.192
female	13189	0	1	0.486	0.500
white	13189	0	1	0.446	0.497

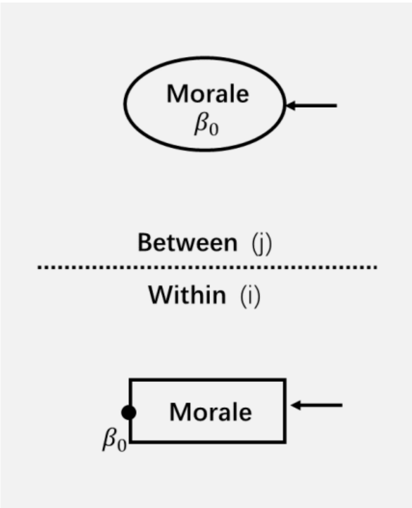
图 9 实例中的描述性信息表

如同前一部分所述，分为 4 个步骤进行分析，每个部分的 Mplus 代码分别如图 10、图 12 图 14、图 16 所示。

3.1 步骤一：将结果变量 MORALE 的方差划分为组内与组间

在一份 Mplus 语句中，相对通用的代码包括 TITLE、DATA、VARIABLE、OUTPUT 这几个部分。TITLE 为代码的标题，于整个程序并无影响，类似于给程序编写者自己的注释。DATA 部分用于指明要分析的数据，本程序中额外写了“Format is”来标注数据矩阵的行列数。VARIABEL 中的“Names”用于给数据中的每一列数值赋予变量名，顺序必须与在数据文件中的变量顺序一致；“UseVariables”标记出本程序中需要使用的变量，只要在这里被定义过的变量才可以被使用，被定义了却没有被使用则会产生 warning；若数据中有缺失值则应该在这一部分中用“Missing”来定义其格式。OUTPUT 指定了希望程序输出的内容，“sampsatat”要求输出样本统计量，“Tech1”要求报告参数设定说明。

在多层模型中特别要注意的语句包括：“Cluster”指出本数据中区分组别的变量，在本例中 depid 变量表示了员工所属的 department 的 ID，这是在多层模型中必然要写的。下面的“BETWEEM=”和“WITHIN=”为特别标记出属于某一层的变量。由于第一步中只有结果变量“morale”一个，而这个变量又将被分为组内和组间两部分，因此不必特别定义在某一层里。ANALYSIS 中要写上“TYPE=TWOLEVEL” 来表示进行的是二层模型的分析（注意，只是随机截距，而没有随机斜率），此时的默认参数估计方法为极大似然估计 ML/MLR。在 MODEL 部分，则是分别写上“%BETWEEN”和“%WITHIN”两句代码来表示组内和组间的模型，如模型图所示对应写上变量以及变量间关系即可。




```
TITLE:      Model 1: Two-level (null) regression model;
DATA:       FILE IS: C:\ Mplus\ch3new.dat;
            Format is 5f8.0,3f8.2;
VARIABLE:   Names are deptid morale satpay female white pctbelow lev1wt
            lev2wt;
            Usevariables are deptid morale;
            Cluster is deptid;
            Between = ;
            Within = ;
ANALYSIS:   TYPE=Twolevel;
Model:      %Between%
            morale;
            %Within%
            morale;
OUTPUT:     Sampstat Tech1;
```

图 10 模型图&Mplus 程序 - 步骤一

如前文所介绍的，在这一步骤中的方程表达为$Y_{ij} = \gamma_{00} + \mu_{0j} + \varepsilon_{ij}$，因此输出的结果将包括截距项$\gamma_{00}$的估计值，以及组内方差$\mathrm{Var}(\varepsilon_{ij})$和组间方差$\mathrm{Var}(\mu_{0j})$，Mplus 将报告估计值、标准误、估计/标准误、显著性（此处将结果整理为两张表格：方差估计、系数估计），如图 11 所示，这三项都是显著不为零的，截距为 26.428，组内方差 33.302，组间方差 5.363，则可以计算组内相关 ICC：$\rho = \sigma_b^2 \,/\, (\sigma_b^2 + \sigma_w^2) = 5.363/(5.363 + 33.302) = 0.139$。

TABLE 3.3 Estimated Variance Components Within and Between Departments

Parameter		*Estimate*	*SE*	*Est./SE*	*Sig.*
Within $\mathrm{Var}(\varepsilon_{ij})$		33.302	0.595	55.995	0.000
Intercept [u_{0j}] $\mathrm{Var}(\mu_{0j})$	Variance	5.363	0.608	8.817	0.000

TABLE 3.4 One-Way ANOVA or "Null Model"

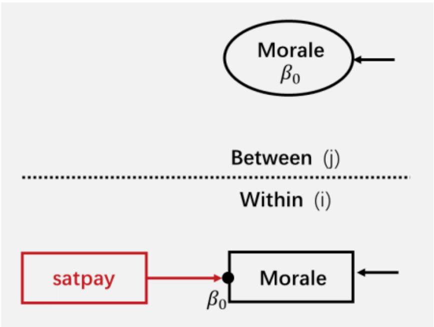
Parameter	*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Intercept γ_{00}	26.428	0.189	139.945	0.000

图 11 步骤一的结果

3.2 步骤二：加入水平 1 的预测变量 SATPAY 和控制变量 FEMALE、WHITE

程序代码方面，相比步骤一共有以下变化：

1. 在 UseVariables 中加入 satpay、female、white 这三个变量；
2. 将这三个变量定义在组内：在 ANALYSIS 中写上"Within=satpay female white"；
3. 为了使斜率在解释时有实际意义，将三个预测变量做中心化处理；
4. 依据模型图在 MODEL 中的"%Within"部分写上，"morale on satpay female white"来表示三个变量对动机的预测作用。




```
TITLE:      Model 2: Level-1 random intercept model;
DATA:       FILE IS: C:\ Mplus\ch3new.dat;
            Format is 5f8.0,3f8.2;
VARIABLE:   Names are deptid morale satpay female white pctbelow lev1wt
            lev2wt;
            Usevariables are deptid morale satpay female white;
            Cluster is deptid;
            Between = ;
            Within = satpay female white;
Define:     Center satpay female white (grand);
ANALYSIS:   TYPE= Twolevel;
Model:      %Between%
            morale;
            %Within%
            morale on satpay female white;
OUTPUT:     SAMPSTAT TECH1;
```

图 12 模型图&Mplus 程序 - 步骤二

此时的方程表达为$Y_{ij} = \gamma_{00} + \mu_{0j} + \beta_{1j}satpay_{ij} + \beta_{2j}female_{ij} + \beta_{3j}white_{ij} + \varepsilon_{ij}$，因此输出的结果除了步骤一的截距项$\gamma_{00}$的估计值，以及组内方差$\mathrm{Var}(\varepsilon_{ij})$和组间方差$\mathrm{Var}(\mu_{0j})$外，还增加了这三个层 1 变量效应的系数估计值$\beta_{1j}$、$\beta_{2j}$、$\beta_{3j}$。首先，由图 13 可以直观的看出，新加入的这三个水平 1 的变量中，satpay 和 white 的效应是显著的(ps<.001)，系数分别为 1.201 和 0.916，而 female 的效应是不显著的(p=.991)。截距与步骤一中基本一致（26.428→26.430），组内方差和组间方差由于被解释了一部分因此都变小了（33.302→17.544, 5.363→1.851）。

TABLE 3.5 Level-1 Random-Intercept Model

Parameter	*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Intercept	26.430	0.114	232.784	0.000
satpay	1.201	0.014	85.725	0.000
female	0.001	0.063	0.011	0.991
white	0.916	0.082	11.181	0.000

TABLE 3.6 Variance Components

Estimates of Covariance Parameters

Parameter		*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Residual		17.544	0.288	60.852	.000
Intercept [u_{0j}]	Variance	1.851	0.231	8.029	.000

图 13 步骤二的结果

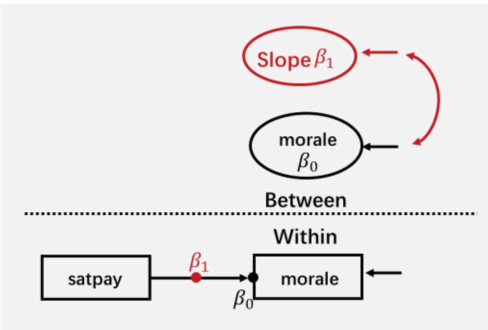
3.3 步骤三：将 SATPAY 对 MORALE 的效应β_{1j}设置为随机项

程序代码方面，相比步骤二共有以下变化：

1. ANALYSIS 部分，TYPE=TWOLEVEL 变为 TYPE=TWOLEVEL RANDOM，变为随机斜率；

2. MODEL 部分

- “%BETWEEN”部分：除了“morale”外，加入“S”表示随机的斜率；加入语句“S with morale”，设定“morale”与“S”相关
- “%WITHIN”部分：“S | morale on satpay”表示将 satpay→morale 的效应随机化




```
TITLE:      Model 3: Specifying a random slope;
DATA:       FILE IS: C:\ Mplus\ch3new.dat;
            Format is 5f8.0,3f8.2;
VARIABLE:   Names are deptid morale satpay female white pctbelow lev1wt
            lev2wt;
            Usevariables are deptid morale satpay female white;
            Cluster is deptid;
            Between = ;
            Within = satpay female white;
Define:     Center satpay female white (grand);
ANALYSIS:   Type = Twolevel random;
Model:      %Between%
            morale S;
            S with morale;
            %Within%
            morale on female white;
            S | morale on satpay;
OUTPUT:     SAMPSTAT TECH1;
```

图 14 模型图&Mplus 程序 - 步骤三

此时的方程为$Y_{ij} = \gamma_{00} + \mu_{0j} + (\gamma_{10} + \mu_{1j})satpay_{ij} + \beta_{2j}female_{ij} + \beta_{3j}white_{ij} + \varepsilon_{ij}$，因此输出的结果与步骤二类似，同样包括截距、三个层 1 效应的系数（但β_{1j}变为γ_{10}）、组内方差、组间方差，增加的分别为：斜率的方差$\mathrm{Var}(\mu_{1j})$、斜率与组间变异的协方差$\mathrm{Cov}(\mu_{0j}, \mu_{1j})$。结果见图 15，在下方的方差估计表格可以看出，斜率的方差为 0.008 虽然小但显著(p=.007)，协方差则是不显著的(p=.734)。

TABLE 3.7 Level-1 Random Intercept and Slope Model

Parameter	*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Intercept	26.436	0.115	229.317	0.000
satpay	1.197	0.014	85.873	0.000
female	0.003	0.063	0.052	0.959
white	0.914	0.082	11.186	0.000

TABLE 3.8 Variance Components

Estimates of Covariance Parameters					
Parameter		*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Residual		17.455	0.292	59.829	0.000
Intercept + slope	Morale	1.861	0.230	8.082	0.000
	Covariance	0.006	0.019	0.340	0.734
	Satpay	0.008	0.003	2.713	0.007

图 15 步骤三的结果

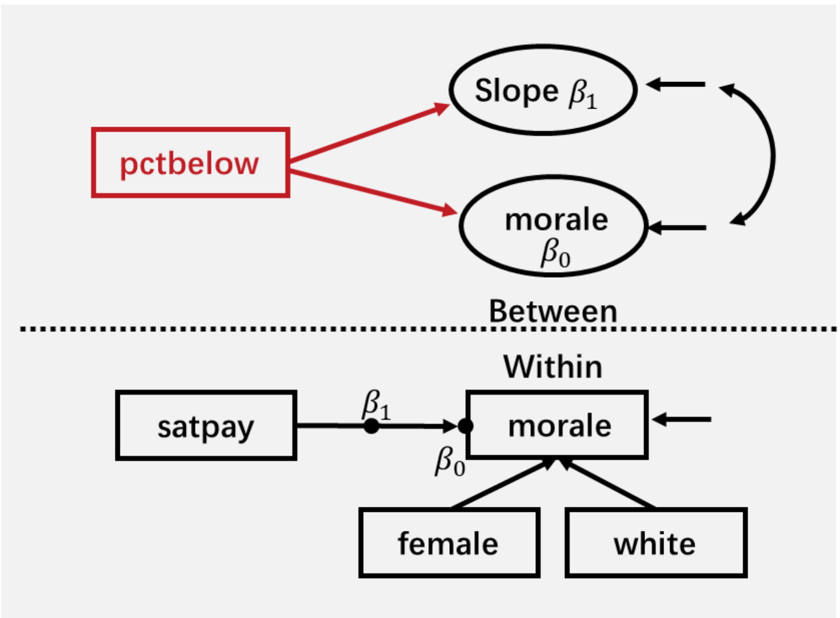
3.4 步骤四：加入水平 2 的预测变量 PCTBELOW，同时指向截距和斜率

程序代码方面（与步骤二有些类似），相比步骤三共有以下变化：

1. 在 UseVariables 中加入 pctbelow 这个变量；

2. 将这个变量定义在组内：在 ANALYSIS 中写上“Between=pctbelow”；

3. 依据模型图在 MODEL 中的“%Between”部分写上，“morale S on pctbelow”来表示组间变量的两个预测作用



```
TITLE:      Model 4: Explaining variation in the level-2 intercept and slope;
DATA:       FILE IS: C:\ Mplus\ch3new.dat;
            Format is 5f8.0, 3f8.2;
VARIABLE:   Names are deptid morale satpay female white pctbelow lev1wt
            lev2wt;
            Usevariables are deptid morale satpay female white pctbelow;
            Cluster is deptid;
            Between = pctbelow;
            Within = satpay female white;
Define:     Center satpay female white pctbelow (grandmean);
ANALYSIS:   Twolevel random;
Model:      %Between%
            morale S on pctbelow;
            S with morale;
            %Within%
            morale on female white;
            S | Morale on satpay;
OUTPUT:     SAMPSTAT TECH1;
```

图 16 模型图&Mplus 程序 - 步骤四

此时的方程为$Y_{ij} = \gamma_{00} + \gamma_{01}pctbelow_j + \mu_{0j} + \left(\gamma_{10} + \gamma_{11}\,pctbelow_j + \mu_{1j}\right)satpay_{ij} + \beta_{2j}female_{ij} + \beta_{3j}white_{ij} + \varepsilon_{ij}$，输出的结果包括截距、三个层 1 效应的系数、两个新加入的层 2 效应的系数（其中有一个本质为跨层级的交互作用）、组内方差、组间方差，斜率的方差、斜率与组间变异的协方差。结果见图 17、图 18，可以看出，两个层 2 的效应只有一个显著：pctbelow→morale 的效应为-0.026，*p*<.001；pctbelow→slope (即 satpay*pctbelow)的效应为 0.001，*p*=.157。

TABLE 3.9 Final Intercept- and Slopes-as-Outcomes Model

Estimates of Fixed Effects				
Parameter	*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Morale Intercept	26.362	0.113	233.072	0.000
pctbelow	–0.026	0.007	–3.574	0.000
female	0.005	0.063	0.074	0.941
white	0.910	0.082	11.064	0.000
Satpay Intercept	1.196	0.014	86.310	0.000
satpay *pctbelow	0.001	0.001	1.416	0.157

图 17 步骤四的结果 - 系数估计

TABLE 3.10 Final Intercept- and Slopes-as-Outcomes Model

Estimates of Covariance Parameters					
Parameter		*Estimate*	*S.E.*	*Est./SE*	*Sig.*
Residual		17.456	0.292	59.789	0.000
Intercept + slope	Morale	1.701	0.218	7.810	0.000
	Covariance	0.014	0.018	0.776	0.438
	Satpay	0.007	0.003	2.455	0.014

图 18 步骤四的结果 - 方差估计

至此一个完整的多层模型分析所包含的四个步骤，以及对应的代数方程、模型图、Mplus 代码、结果解释都已经介绍完毕。此时，通过第四个步骤得到的最终结果，我们可以验证开始提出的两个假设：

1. 员工对于薪水的满意程度将显著的正向预测他们的工作斗志(1.196, *p*<.001)，控制变量性别的效应不显著(0.005, *p*=.941)、种族的效应显著(0.910, *p*<.001)

2. 所处部门的整体工资水平也能显著预测他们的工作斗志（-0.026, *p*<.001），但并不能在薪水满意度和斗志的关系间起调节作用(0.001, *p*=.157)

此外，我们还可以进行一个分析，比较步骤一与步骤四的两个模型中，组间误差方差和组内误差方差分别减少了百分之多少，类似于回归方程中的R^2，可以看出这个理论模型解释了结果变量方差的比例：组内方差变化$= (33.302 - 17.456)/33.302 = 0.476\ (47.6\%)$，组间方差变化$= (5.363 - 1.701)/5.363 = 0.683\ (68.3\%)$。