

# 纳入式分类分析法在潜在剖面模型的 后续多元回归中的应用\*

张洁婷<sup>1</sup>, 张敏强<sup>2</sup>, 焦 璨<sup>1</sup>

(1. 深圳大学心理与社会学院, 深圳 518060; 2. 华南师范大学心理学院, 心理应用研究中心, 广州 510631)

**摘 要:** 纳入式分类分析法能克服传统的分类分析法对后续一元回归模型参数的低估, 发挥潜在类别模型的后续分析简化变量间交互作用的功能。本文进一步将纳入式分类分析法拓展至潜在剖面模型后续的多元统计分析中。通过蒙特卡洛模拟实验, 比较各种纳入变量的方法思路与后续分析模型在四种常见的多元回归模型中参数估计的表现。结果发现, 纳入式分类分析法所需纳入的变量取决于后续分析中与因变量、潜类别变量的关系, 且后续分析使用含交互作用的模型更为稳健。

**关键词:** 潜在剖面模型; 后续分析; 纳入式分类分析法; 多元回归; 中介; 调节

中图分类号: B841.2

文献标识码: A

文章编号: 1003-5184(2019)01-0040-07

## 1 引言

潜在剖面模型(latent profile model, LPM)是潜在类别模型(latent class model, LCM)的拓展模型, 主要处理连续型外显变量的潜在分类, 更适用于心理测量常见的结果数据。该方法能反映类别之间在每个维度上的量化差异, 使得心理学研究兼有质化研究的全面、多维性以及量化研究的尺度性(Marsh, Lüdtke, Trautwein, & Morin, 2009; 马文超, 边玉芳, 骆方, 2012)。

心理学与行为学的交叉研究一般同时出现连续型的量表数据与离散型的外显行为数据, 这时可以使用潜类别变量作为“桥梁”对两者的关系进行建模, 简化多个变量之间较为繁杂的交互作用(Lanza & Rhoades, 2013)。譬如, 在探讨人格对家庭教养与品行问题关系的调节作用时, 传统的多元线性回归需考虑多项人格维度作为调节变量, 且这些调节变量之间也很可能存在交互作用。使用大五人格测验则存在  $C_2^2 + C_3^3 + C_4^4 + C_5^5 = 26$  种交互作用项, 若要进一步分析维度间的交互作用, 将各维度分成高低两水平, 则有种组合的人格特征, 这大大增加了计算量和统计检验, 从而也会削弱统计功效(Lanza & Rhoades, 2013; Merz & Roesch, 2011)。而且在实际研究中, 并非每种维度组合都有实质性意义, 大多集中在固定的几类人格(De Clercq, Rettew, Althoff, &

De Bolle, 2012; Kinnunen et al., 2012)。针对该实证研究, 可以使用 LPA 来剖析出几种主要的人格类型, 并将分类变量作为后续模型的变量, 与其他自变量或因变量(统称为附属变量, auxiliary variable)进一步建模, 这能更有效地处理人格维度之间的高阶交互效应项。这种统计分析可统称为后续分析(subsequent analysis, Peterson, Bandeen - Roche, Budtz - Jørgensen, & Larsen, 2012), 尤其适用于复杂的中介或调节模型中(Donnellan & Robins, 2010; Martel, Goth - Owens, Martinez - Torteya, & Nigg, 2010), 有助于降低接下来推断统计的一类错误率(Lanza & Rhoades, 2013)。

后续分析拓宽了类别型潜变量在结构模型中的应用, 也常见于 LCA/LPA 的实证研究中(如, 杨丽珠, 马世超, 2014; 杨晶晶, Lowenstein, Jackson, 郑涌, 2013)。然而, 大部分的后续分析所使用的传统的三步法, 即直接将潜在类别作为确定的分类进入下一步的分析, 会低估潜类别变量与其他变量之间的关系程度。针对此方法问题, 陆续提出了各种后续分析方法, 包括一步法(Clark & Muthén, 2009; Vermunt, 2010)、基于模型方法(Lanza, Tan, & Bray, 2013)、Bartlett 分数法(Peterson, Bandeen - Roche, Budtz - Jørgensen, & Larsen, 2012)、ML 三步法和 BCH 法(Vermunt, 2010)。其中, BCH 和 ML 方法在

\* 基金项目: 国家自然科学基金资助项目(31700982), 广东省哲学社会科学“十三五”规划项目(GD17YX101), 深圳市教育科学“十三五”规划招标课题(zdzz16004)。

通讯作者: 张敏强, E-mail: 2640726401@qq.com。

模拟实验中得到较为理想的估计精确度,但在实际应用中常受到模型复杂性、边界值等计算问题所限制(Asparouhov & Muthén, 2014; Bakk, Tekle, & Vermunt, 2013; Bray, Lanza, & Tan, 2015; Lanza, Tan, & Bray, 2013; 张洁婷, 张敏强, 黎光明, 2017)。BCH方法虽然是这些方法中最适合用于中介或调节模型中,但仅适于因变量为连续变量的模型(Asparouhov & Muthén, 2015)。

Bray, Lanza 和 Tan(2015)提出了纳入式分类分析法(Inclusive Classify - analyze, 本文简称纳入法),并通过模拟实验证明纳入式分类分析法能提高后续分析模型的参数估计精度。张洁婷等(2017)模拟研究发现,该方法的估计精度相比于传统方法有了很大的提升,仅略低于 BCH 或 ML 三步法。而且,该方法能处理因变量为二分变量的后续分析,应用灵活简便的优点,具有方法推广的潜力与价值。进一步的模拟实验发现,在针对一元回归后续分析模型的各种条件下,纳入法的参数估计结果最为稳健。该方法借助多重计算方法(multiple imputation)用于处理缺失值的统计思想(Collins, Schafer, & Kam, 2001):在对缺失值进行多重计算时,充分纳入与其随机缺失机制有关的附属变量信息,能提高缺失值的估计有效性并降低估计偏差。潜类别变量可看作一个完全缺失的变量,将(后续)分析模型中所涉及的变量  $y$  作为协变量进入 LCM/LPM(即测量模型)中(如公式 1),使测量模型所得的后验概率包含了分析模型中涉及的所有变量与潜类别变量之间关系的先验信息(如公式 2),因而测量模型和分析模型有同等全面的信息,以修正对潜类别变量与因变量之间关系的低估。

$$f(i|Y=y) = \sum_{t=1}^T P(C=t|Y) \prod_{j=1}^J f(i_j|\mu_{jt}, \sigma_{jt}^2) \quad (1)$$

$$P(C=t|I_{jk}, y) = \frac{P(C=t|y)f(I_{jk}|C=t, y)}{f(I_{jk}|y)} \quad (2)$$

纳入式分类分析法相比于其他方法操作简单直观,仅需在第一步的测量模型中增加协变量,能适应更复杂的多元回归后续分析。但现阶段该方法的应用仅限于在简单的潜类别变量与因变量简单的一元回归关系。对于复杂的多元回归情形,还需进一步探究。根据 Bray 等(2015)纳入法思想:后续分析的模型中与潜类别变量有关的所有变量都应纳入潜在类别模型中,以免低估潜类别变量与这些变量之间的关系。基于此方法思想,本研究提出在各种后续

多元回归分析中纳入法的应用思路,并通过模拟研究在常用的几种多元回归数据模型中进行验证。

## 2 方法

### 2.1 数据模型及产生

本模拟实验的真值模型如图 1。

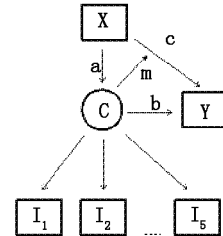


图1 LPM 测量模型及后续统计模型

在测量模型中,从简版大五人格的潜在剖面分析实证研究中抽取常见的三种人格类型(Kinnunen et al., 2012; Rammstedt, Riemann, Angleitner, & Borkenau, 2004),设由 5 项外显变量组成的三类别潜在剖面模型,类别概率均匀分布(即都是 1/3);各类外显指标的条件均值如表 1,各外显变量的整体标准差均设为 0.5,分类区分度良好(熵=0.8)。

表1 外显指标参数表

	C1	C2	C3
I1	1	0	-1
I2	-1	0	1
I3	-1	0	1
I4	-1	0	1
I5	-1	0	1

在后续分析中,有另一个自变量  $X$  和一个因变量  $Y$ ,  $X$  和  $Y$  都是二分变量,基于 LPM 估计的最大后验概率得到分类变量  $W$ ,分析模型方程如下:

$$\begin{aligned} \text{LogOdds}(Y|W=w, X=x) &= \\ \ln \frac{P(Y=1|W=w, X=x)}{P(Y=0|W=w, X=x)} &= \beta_0 + \beta_x x + \\ &\beta_{w_1} w_1 + \beta_{w_3} w_3 + \beta_{m_1} m_1 + \beta_{m_3} m_3 \end{aligned} \quad (3)$$

以  $C=2, X=0$  为参考系,其中  $w_1$  和  $w_3$ , 表示潜类别 1 和 3,  $m_1 = xw_1$ ,  $m_3 = xw_3$ 。

有两项实验因素:(1)  $X$  对  $C$  的影响( $a$ ):无(0),有(Cohen 效果量  $\omega=.3$ );(2)交互作用( $m$ ):无( $\beta_{m_1}=0, \beta_{m_3}=0$ ),有( $\beta_{m_1}=2, \beta_{m_3}=-2$ ),组合形成四种常用的回归模型,如图 2。其他回归系数设为:截距项  $\beta_0=-2$ ,潜类别变量  $C$  对因变量  $Y$  的影响(路径  $b$ )  $\beta_{w_1}=1, \beta_{w_3}=0.5$ ,预测变量  $X$  对  $Y$  的影响(路径  $c$ )  $\beta_x=1$ ,每批数据的样本量设为  $N=3000$ 。

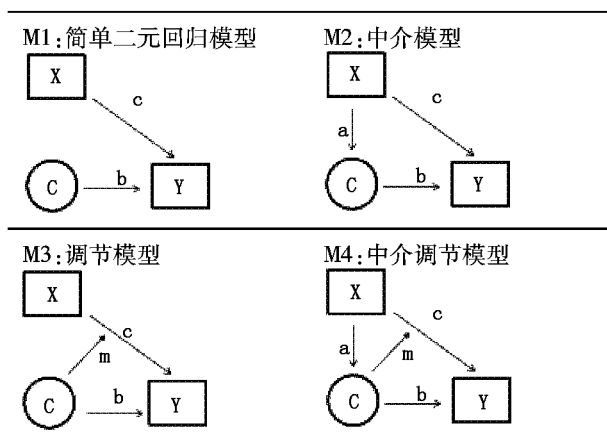


图2 四种后续分析统计模型\*

## 2.2 方法比较

在该潜在剖面模型的后续分析中,分类分析法

表2 回归模型 A 与 B 系数与  $\text{LogOdds}(Y|C=w, X=x)$  对应表

回归系数		$\text{LogOdds}$
$\beta_0$	$\gamma_0$	$\text{LogOdds}(Y C=2, X=0)$
$\beta_{w_1}$	$\gamma_{w_1}$	$\text{LogOdds}(Y C=1, X=0) - \text{LogOdds}(Y C=2, X=0)$
$\beta_{w_3}$	$\gamma_{w_3}$	$\text{LogOdds}(Y C=3, X=0) - \text{LogOdds}(Y C=2, X=0)$
$\beta_x$	$\gamma_x$	$\text{LogOdds}(Y C=2, X=1) - \text{LogOdds}(Y C=2, X=0)$
$\beta_{w_1} + \beta_{m_1}$	$\gamma_{w_1}$	$\text{LogOdds}(Y C=1, X=1) - \text{LogOdds}(Y C=2, X=1)$
$\beta_{w_3} + \beta_{m_3}$	$\gamma_{w_3}$	$\text{LogOdds}(Y C=3, X=1) - \text{LogOdds}(Y C=2, X=1)$

## 2.3 实验程序

首先,对以上四种模型条件分别使用蒙特卡洛模拟产生 500 批数据。然后,对每批模拟数据分别建立五种 LPM,即传统的 LPM(不带协变量)与四种纳入式 LPM,类别之间外显变量的标准差设为相等。使用 100 个随机初始值(random set)进行极大似然估计,以确保模型得到识别。使用最优归类法分别对基于五种 LPA 模型所得到的后验概率计算最优分类变量(W)。最后,基于五种 LPM 的分类值 W,分别建立 logistic 回归模型,其中所选取的回归模型分为去交互作用和含交互作用的模型,估计得到回归系数。以上模拟与分析主要使用软件 R3.1.0 和 Latentgold4.5 实现,首先以对数发生比 Logodds 作为初步的模型拟合度评价指标,再以平均

偏差  $\frac{\sum (\hat{\theta}_i - \theta)}{N}$  与均方根误差 (RMSE,  $\sqrt{\frac{\sum (\hat{\theta}_i - \theta)^2}{N}}$ ) 作为比较指标,比较 10 种分析方法的参数估计,推断各真值模型下的最优方法以及方法改进的规律。

在使用中需要考虑以下两个因素,即形成  $5 \times 2$  种拟使用的方法。

(1) LPM 模型纳入的项目:传统分类分析法(简称传统法);纳入 Y(简称纳入 Y 法);纳入 X(简称纳入 X 法);纳入 X 和 Y(简称纳入 XY 法);纳入 X, Y 与乘积项(简称全纳入法);

(2) 后续分析模型是否含交互作用:含交互作用(模型 A,如式 3),去交互作用(模型 B,  $=0, m_2=0$ ,如式 4)。

$$\text{LogOdds}(Y|C=w, X=x) = \beta_0 + \beta_x x + \beta_{w_1} w_1 + \beta_{w_3} w_3 \quad (4)$$

各系数在两个模型中所代表的意义有所不同,因此模型 B 的回归系数用  $\gamma$  相应替换,具体如表 2。

## 3 结果

各模型下模型对数发生比与回归系数的估计平均偏差、RMSE 见附录表 1~4。在模型 M1(简单的二元回归模型)和模型 M2(中介模型)中,预测变量 X 和潜类别变量 C 对因变量 Y 无交互作用。各 LPM 模型下,分析模型 A(含交互作用)和 B(不含交互作用)在 Logodds 和回归系数(模型 A 的回归系数  $\beta_0, \beta_{w_1}, \beta_{w_3}, \beta_x$  与模型 B 中的  $\gamma_0, \gamma_{w_1}, \gamma_{w_3}, \gamma_x$ )上呈现相似的平均偏差,只是模型 A 各参数的 RMSE 稍高于模型 B。这说明两种分析模型在不含交互作用的真值模型中具有等价性,且都具有较高的精确性。在模型 M3(调节模型)和模型 M4(中介调节模型)中,若使用不含交互作用的后续分析模型 B,无论使用哪种纳入法都无法弥补模型 B 忽略交互作用所导致对数发生比的明显偏差。这说明,当真值模型不含交互作用时,在统计模型中加入或不加入交互作用项对模型拟合结果相差不大;而当真值模型含交互作用时,需要在统计模型中加入交互作用项,否则拟合度严重受影响。基于此结果,进

\* C:三类别的潜在类别变量;Y:因变量(二分变量);I:五个外显变量(连续变量);X:预测变量(二分变量)

一步比较各测量模型下后续回归模型的参数估计偏差。

各 LPM 测量模型下回归系数估计的 RMSE 相差不大,但平均偏差具有更明显的差异。如图 3,对于 M1 和 M2,各 LPA 方法在  $\beta_x$ ,  $\beta_{m_1}$  和  $\beta_{m_3}$  的估计上都较为接近真值,各 LPA 方法对回归系数的估计差异主要集中在  $\beta_0$ ,  $\beta_{w_1}$ ,  $\beta_{w_3}$  (或  $\gamma_0$ ,  $\gamma_{w_1}$ ,  $\gamma_{w_3}$ )。在 M1,传统 LPA 方法对这些回归系数估计产生最大的低估偏差,只纳入预测变量  $X$  对参数没有任何改进,同样产生最大的低估。纳入因变量  $Y$  则大大克服了低估现象,且偏差最小;其他方法在纳入  $Y$  的基

础上纳入  $X$ 、乘积项,未进一步降低参数估计的偏差,反而有轻微增加。对于 M2,同样地,传统分类分析法与纳入  $X$  法的估计偏差依然最大,同时,纳入  $X$   $Y$  法的偏差、RMSE 小于只纳入  $Y$  的方法,但与全纳入法一样,相比纳入  $Y$  法精确度提高的幅度较轻微。对于 M3、M4 模型,传统分类分析法与纳入  $X$  方法的估计偏差最大,全纳入法的偏差最小,其他纳入法居中。虽然纳入  $Y$  法和纳入  $XY$  法对  $\beta_0$ ,  $\beta_{w_1}$ ,  $\beta_{w_3}$  和  $\beta_x$  的低估有所缓和,但在交互作用系数  $\beta_{m_1}$  和  $\beta_{m_3}$  的估计依然存在很大的低估。完全纳入法存在低于 0.1 的高估,偏差相对最小。

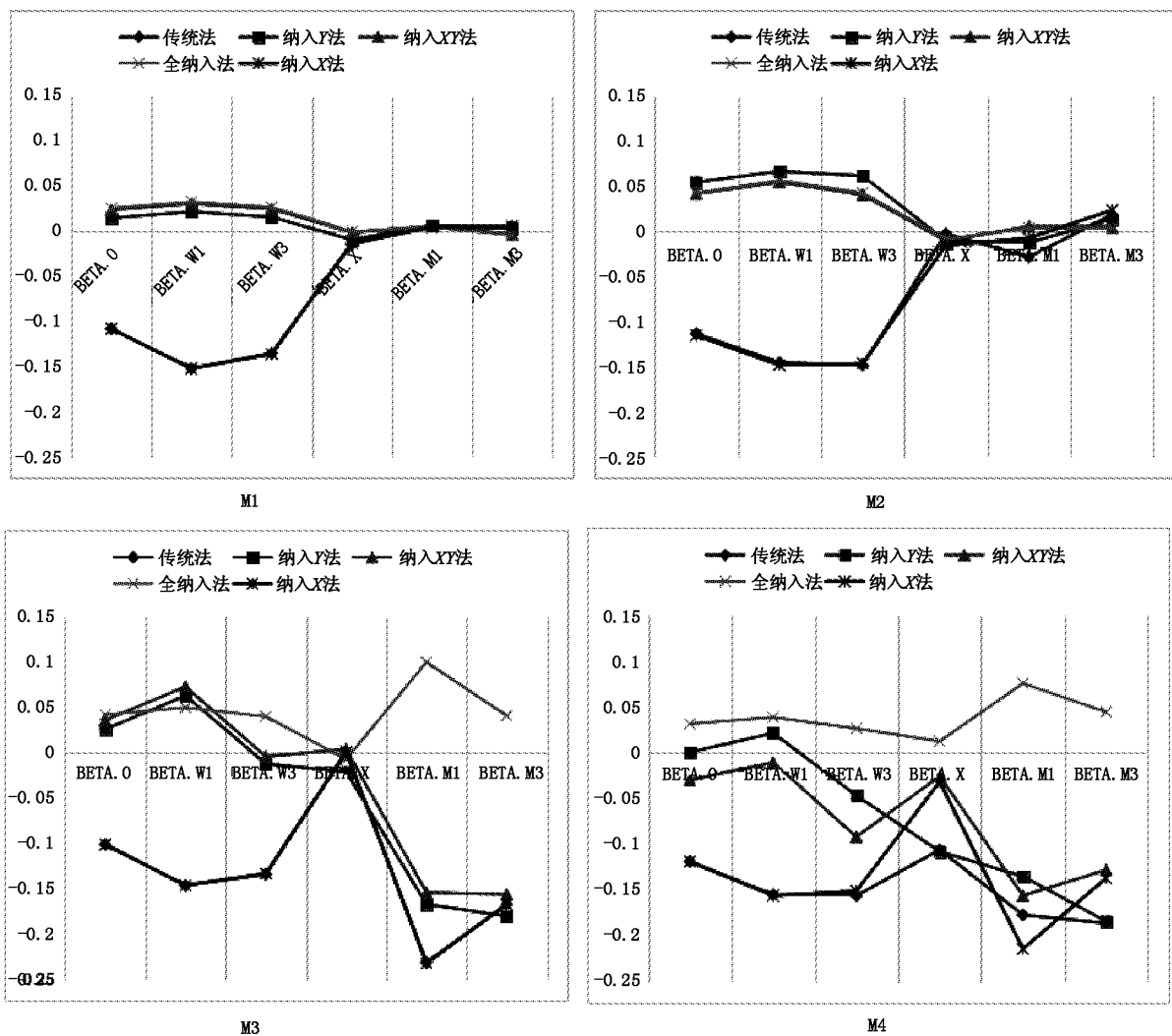


图 3 四种模型下回归系数的估计偏差

## 4 讨论

### 4.1 纳入式分类分析法在多元回归模型中的应用

通过比较各方法在参数估计的精确度,分析纳

入式分类分析法在四种常见多元回归模型下具体所需纳入的变量及其后续统计模型的选用。各真值模型下的方法表现总结如表 3。

表3 模拟实验结果总结表

	模拟实验设计		结果	
	X对C的影响	有无交互作用	最优测量模型	最优后续分析模型
M1:简单二元回归	无	无	纳入Y	两种皆可
M2:中介模型	有	无	纳入X和Y	两种皆可
M3:调节模型	无	有	纳入X,Y,X×Y	含交互作用的模型
M4:中介调节模型	有	有	纳入X,Y,X×Y	含交互作用的模型

对于LPM测量模型的选用:(1)纳入因变量Y的方法能提高因变量对潜类别变量的回归参数估计精度,这与前人研究结果一致;同时,也发现若只纳入预测变量X,无法提高Y对C回归系数的估计精度,与传统方法有同样大的估计偏差。且当预测变量X对潜类别变量C有影响时,在纳入Y的基础上纳入X仅小幅度地提高了精确度。(2)当预测变量与潜类别变量对因变量存在交互作用时,全纳入法,即在纳入X和Y的同时纳入乘积项 $X \times Y$ 能提高参数估计精度,尤其对交互作用系数的估计中起到必要作用(是纳入Y法偏差的 $1/4 \sim 3/5$ )。本模拟研究表明,在后续多元回归中,不但需要纳入与潜类别变量直接关系的因变量,还需纳入涉及交互作用的变量及其乘积项,即验证了Bray等(2015)提出的理论假设。

对于后续分析模型的选用:对于不含交互作用的真值模型,使用含交互作用的分析模型和去交互作用的模型在回归参数估计和对数发生比上结果一致。而在含交互作用的真值模型中,使用去交互作用的模型,会产生很大的偏差,且所有的分类分析法都无法弥补这种偏差。这反映了带交互作用的模型对实际数据拟合的稳健性:即使在真值模型中不存

在交互作用,在测量模型或后续分析中考虑交互作用都不会对后续分析中的交互作用产生明显的估计偏差。

#### 4.2 纳入式分类分析法在后续多元回归分析中的应用推广

使用纳入式分类分析法处理后续多元回归模型时,可根据理论假设考虑是否纳入乘积项;对于调节作用在较含糊或没有足够理论先验信息的情况下,设想可以通过以下步骤(如图4)逐一将备选调节变量进入测量与后续统计模型,以在避免明显偏差的前提下寻找合理的调节变量。首先初步纳入因变量Y建立测量模型,然后通过比较含交互作用的模型与去交互作用的模型的拟合度差异,进而检验交互作用的显著性。若交互作用显著,建议使用全纳入法(即纳入X,Y和乘积项)重新建立测量模型,最后重新建立含交互作用的分析模型,以确保各参数估计的精确性。若交互作用不显著,进一步分析预测变量X对潜类别变量C的影响:若潜类别变量C受预测变量X的影响,重新使用纳入XY法建立测量模型,再进一步估计去交互作用模型的参数;若潜类别变量C不受预测变量X的影响,可直接使用去交互作用的纳入Y法。

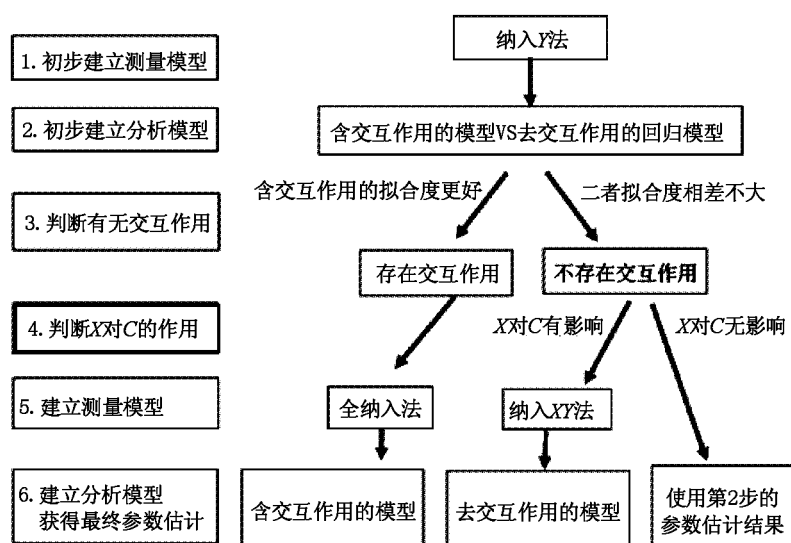


图4 后续的多元回归分析步骤

### 4.3 研究展望

本研究对 Bray 等(2015)所提出的纳入式分类分析法思想进行拓展研究,现阶段仍有许多进一步探讨的议题。其一,当纳入过多变量时很可能与一步法相似地发生类别转移或类别意义改变的情况,对于探索性的研究目的,或许可以先探讨附属变量与潜在类别外显指标的关系矩阵,或比较纳入和不纳入相应变量的分类一致性,以初步判断附属变量与潜类别变量之间的关系强度(Huang, Brecht, Hara, & Hser, 2010)。该构想还有待模拟验证。其二,仅以 logistic 回归模型为后续分析模型,对于线性回归后续分析(Bakk, Oberski, & Vermunt, 2015; Bakk & Vermunt, 2015)、带调节的中介或带中介的调节模型(温忠麟,叶宝娟,2014)等更复杂的后续关系并未得以研究。此外,对于模型的效果量估计、统计功效等回归分析的重要参考指标,还有待软件的开发实现,进一步作为方法比较的参考依据。

### 5 结论

对于潜在剖面模型的后续分析,纳入式分类分析法所需纳入的变量取决于后续分析中自变量  $X$  与因变量  $Y$ 、潜类别变量  $C$  的关系。在所构建的后续回归模型中,当  $C$  与  $Y$  存在关系,有必要纳入  $Y$ ; 当  $X$  与  $C$  存在关系,纳入  $X$  对参数估计偏差有轻微的改善;当  $C$  与  $X$  交互影响  $Y$ ,则有必要纳入  $X$  与  $Y$  及其乘积项,并且需要在后续统计模型中包含交互作用项。在探索性的建模分析中可以依次对交互作用、 $X$  与  $C$  的关系进行检验,从而选择合适的 LPM 测量模型与后续统计模型。

### 参考文献

- 马文超,边玉芳,骆方.(2012).网络成瘾的潜在结构:连续的还是分类的?《心理发展与教育》,5,554-560.
- 温忠麟,叶宝娟.(2014).有调节的中介模型检验方法:竞争还是替补?《心理学报》,45(9),714-726.
- 杨晶晶,Lowenstein, A., Jackson, T., 郑涌.(2013).代际团结潜在类别与关系质量对自陈健康及幸福感的影响.《心理学报》,7,811-824.
- 杨丽珠,马世超.(2014).初中生人格类型划分及人格类型发展特点研究.《心理科学》,37(6),1377-1384.
- 张洁婷,张敏强,黎光明.(2017).潜在剖面模型的后续分析——比较分类分析法改进后的偏差.《心理学探新》,37(5),434-440.
- Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture Modeling: Three - Step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 329-341.
- Asparouhov, T., & Muthén, B. (2015). Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary second model. *Mplus Web Notes*, (21), 1-22.
- Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias - adjusted three - step approaches. *Sociological Methodology*, 43, 272-311.
- Bakk, Z., Oberski, D., & Vermunt, J. K. (2016). Relating latent class membership to continuous distal outcomes: Improving the LTB approach and a modified three - step implementation. *Structural Equation Modeling*, 23(2), 278-289.
- Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling*, 23(1), 20-31.
- Bray, B. C., Lanza, S. T., & Tan, X. (2015). Eliminating bias in classify - analyze approaches for latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(1), 1-11.
- Clark, S. L., & Muthén, B. (2009). *Relating Latent Class Analysis Results to Variables not Included in the Analysis*. Retrieved from <https://www.statmodel.com/download/relatinglca.pdf>
- Collins, L., Schafer, J., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6(4), 330-351.
- De Clercq, B., Rettew, D., Althoff, R. R., & De Bolle, M. (2012). Childhood personality types: Vulnerability and adaptation over time. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 53(6), 716-722.
- Donnellan, M. B., & Robins, R. W. (2010). Resilient, overcontrolled, and undercontrolled personality types: Issues and controversies. *Social and Personality Psychology Compass*, 4(11), 1070-1083.
- Huang, D., Brecht, M., - L., Hara, M., & Hser, Y. - I. (2010). Influences of a covariate on growth mixture modeling. *Journal of Drug Issues*, 40(1), 173-194.
- Kinnunen, M. - L., Metsäpelto, R. - L., Feldt, T., Kokko, K., Tolvanen, A., Kinnunen, U., ... Pulkkinen, L. (2012). Personality profiles and health: Longitudinal evidence among Finnish adults. *Scandinavian Journal of Psychology*, 53(6), 512-522.
- Lanza, S. T., & Rhoades, B. L. (2013). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science*, 14, 157-168.
- Lanza, S. T., Tan, X., & Bray, B. C. (2013). Latent class analysis with distal outcomes: A flexible model - based approach. *Structural Equation Modeling: A Multidisciplinary Journal*, 20

- (1), 1 – 26.
- Marsh, H. W. , Lüdtke, O. , Trautwein, U. , & Morin, A. J. S. (2009). Classical latent profile analysis of academic self – concept dimensions; Synergy of person – and variable – centered approaches to theoretical models of self – concept. *Structural Equation Modeling: A Multidisciplinary Journal*, 16, 191 – 225.
- Martel, M. M. , Goth – Owens, T. , Martinez – Torteya, C. , & Nigg, J. T. (2010). A person – centered personality approach to heterogeneity in attention – deficit/ hyperactivity disorder (ADHD). *Journal of Abnormal Psychology*, 119, 186 – 196.
- Merz, E. L. , & Roesch, S. C. (2011). A latent profile analysis of the Five Factor Model of personality: Modeling trait interactions. *Personality and Individual Differences*, 51 (8) , 915 – 919.
- Peterson, J. , Bandeen – Roche, K. , Budtz – Jørgensen, E. , & Larsen, K. G. (2012). Predicting latent class scores for subsequent analysis. *Psychometrika*, 77(2) , 244 – 262.
- Rammstedt, B. , Riemann, R. , Angleitner, A. , & Borkenau, P. (2004). Resilients, Overcontrollers, and Undercontrollers: The replicability of the three personality prototypes across informants. *European Journal of Personality*, 18(1) , 1 – 14.
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three – step approaches. *Political Analysis*, 18 (4) , 450 – 469.

## Application of Inclusive Classify – analyze Approach for Multiple Regression Subsequent to Latent Profile Model

Zhang Jieting<sup>1</sup>, Zhang Minqiang<sup>2</sup>, Jiao Can<sup>1</sup>

(1. College of Psychology and Sociology, Shenzhen University, Shenzhen 518060;

2. Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou 510631)

**Abstract:** Inclusive classify – analyze can overcome the underestimation of simple regression parameters by the traditional classify – analyze approach in subsequent analysis of latent profile model, and manages to simplify the estimation of interaction among variables. The current study aimed to extend the inclusive approach to some multiple regression analysis subsequent to latent profile model. Monte Carlo simulation study was conducted to investigate which variables should be included in the latent profile model, and whether interaction should be considered in the subsequent analytic model. Data was generated based on regression model of a binary distal outcome to a three – profile variable and a binary predictor. Four commonly – used model scenarios were considered; multiple regression model, and regression models with moderator, mediator or both. The results confirm the necessity of adding distal outcome to the measurement model, and also suggest that, variables related to or interacted with the latent profile variable should be included in the latent profile model. Failure to include these variables would induce attenuation of the estimated effects. In addition, we found model which include interactions term is more robust in the subsequent analysis.

**Key words:** latent profile model; subsequent analysis; inclusive classify – analyze approach; multiple regression; mediation; moderation