

# 基于多元回归的调节效应分析\*

方杰\*\*<sup>1</sup> 温忠麟<sup>2</sup> 梁东梅<sup>2</sup> 李霓霓<sup>3</sup>

(<sup>1</sup> 广东财经大学人文与传播学院, 广州, 510320)

(<sup>2</sup> 华南师范大学心理学院 / 心理应用研究中心, 广州, 510631) (<sup>3</sup> 广州美术学院, 广州, 510260)

**摘要** 在心理学和其他社科研究领域, 大量实证研究建立调节模型, 以分析自变量对因变量关系的影响机制, 但在基于多元回归的调节效应分析实践中仍存在不足。我们回顾了均值中心化在基于多元回归的调节效应分析中的作用, 均值中心化不影响乘积项 (即调节效应) 的检验, 仅对一阶项 (即主效应) 的检验有影响。讨论了简单斜率的检验方法, 建议在调节变量为连续变量时, 使用 Johnson-Neyman 法进行简单斜率检验; 在调节变量为类别变量或研究者对某个调节变量值感兴趣时, 使用选点法。并用一个实际例子演示如何进行调节效应分析。随后展望了调节效应检验的拓展方向。

**关键词** 调节效应 多元线性回归 均值中心化 选点法 Johnson-Neyman 法

## 1 前言

调节 (moderation) 是社会科学研究中重要的方法学概念, 是研究者探索多个变量之间关系的重要手段。如果因变量  $Y$  和自变量  $X$  的关系 (回归斜率的大小和方向) 随第三个变量  $Z$  的变化而变化, 则称  $Z$  在  $X$  和  $Y$  之间起调节作用, 此时称  $Z$  为调节变量 (见图 1(a))。当  $X$  和  $Z$  都可以作为定距变量时, 调节效应可以通过下面回归方程进行分析 (图 1(b) 是相应的路径图),

$$Y = i + aX + bZ + cXZ + \varepsilon \quad (1)$$

如果回归系数  $c$  显著, 就表示调节效应显著。调节效应的效果量用  $\Delta R^2 = R_1^2 - R_0^2$  表示, 其中  $R_1^2$  表示回归方程 (1) 的测定系数,  $R_0^2$  表示去掉乘积项  $XZ$  后的回归方程 (1) 的测定系数 (温忠麟, 刘红云, 侯杰泰, 2012)。

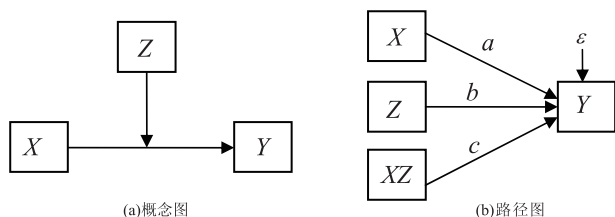


图 1 调节效应模型图

在基于多元回归的调节效应分析中, 研究者通

常将自变量  $X$  和调节变量  $Z$  进行均值中心化 (即变量减去样本均值, 以下简称中心化), 然后再产生乘积项  $XZ$ , 最后进行多元回归分析, 这似乎已成为调节效应分析的标准步骤。但在多元回归的调节效应分析中, 有研究者认为中心化的作用往往被过分夸大了 (Dalal & Zickar, 2012; Echambi & Hess, 2007; Hayes, 2013)。另外, 当调节效应显著时, 研究者常将方程 (1) 重写可得方程 (2),

$$Y = (i + bZ) + (a + cZ)X + \varepsilon \quad (2)$$

其中  $a + cZ$  表示简单斜率 (simple slope), 它反映了因变量  $Y$  和自变量  $X$  的关系是如何受变量  $Z$  调节的。目前, 研究者通常采用选点法 (pick-a-point approach) 进行简单斜率的显著性检验 (Alike & West, 1991; Cohen, Cohen, West, & Alike, 2003), 但近年来, 选点法遭到部分学者的批评 (Bauer & Curran, 2005; Hayes, 2013; Hayes & Matthes, 2009; Jose, 2012; Preacher, Curran, & Bauer, 2006; Spiller, Fitzsimons, Lynch, & McClelland, 2013)。本文的目的是评述多元回归的调节效应分析中, 中心化策略的作用和简单斜率检验的方法, 并根据数据类型向研究者推荐合适的简单斜率检验方法。

## 2 中心化

\* 本研究得到国家自然科学基金项目 (31271116)、教育部人文社会科学研究青年基金项目 (14YJC190003) 和广东省哲学社会科学“十二五”规划项目 (GD13CXL01) 的资助。

\*\* 通讯作者: 方杰。E-mail: fangjie@gdufe.edu.cn

## 2.1 中心化减少非本质的多重共线性

变量中心化后, 分析调节效应的回归方程可以写成

$$Y = i' + a'(X - \bar{X}) + b'(Z - \bar{Z}) + c'(X - \bar{X})(Z - \bar{Z}) + \varepsilon' \quad (3)$$

多数研究者都认为中心化的目的之一是为了减少多重共线性 (Dalal & Zickar, 2012; Jose, 2012)。

这里需要澄清两个问题。第一, 方程 (1) 存在两种多重共线性的情况, 一种是自变量  $X$  和乘积项  $XZ$  之间、调节变量  $Z$  和乘积项  $XZ$  之间存在共线性的情况; 另一种是自变量  $X$  和调节变量  $Z$  之间存在共线性的情况, 那么中心化减少的是哪种共线性情况呢?

如果  $X$  和  $Z$  服从正态分布,  $X$  和  $XZ$  的协方差为 (Bohrstedt & Goldberger, 1969):

$$\text{cov}(X, XZ) = E(Z) \text{var}(X) + E(X) \text{cov}(X, Z) \quad (4)$$

将自变量  $X$  和调节变量  $Z$  按均值进行中心化后, 其数学期望 (即均值)  $E(X)$  和  $E(Z)$  都变为 0, 由公式 (4) 可知自变量  $X$  和乘积项  $XZ$  之间的协方差为 0, 所以自变量  $X$  和乘积项  $XZ$  之间的相关为 0, 因而不存在共线性情况。即使正态假设不满足, 均值中心化也可以减小  $X$  和  $XZ$  共线性的严重性。同理可知, 均值中心化可以减小  $Z$  和  $XZ$  共线性的严重性。因此均值中心化减小的是第一种共线性情况 (温忠麟等, 2012; Alike & West, 1991; Dalal & Zickar, 2012; Fairchild & McQuillin, 2010; Hayes, 2013)。

第二个需要澄清的问题是, 中心化减少多重共线性对调节效应检验结果起怎样的作用呢? 将方程 (3) 进行重写得到方程 (5),

$$Y = (i' - a'\bar{X} - b'\bar{Z} + c'\bar{X}\bar{Z}) + (a' - c'\bar{Z})X + (b' - c'\bar{X})Z + c'XZ + \varepsilon' \quad (5)$$

将方程 (5) 和 (1) 比较可知, 中心化前后, 乘积项  $XZ$  的回归系数没有变化 (即  $c = c'$ ) (Dalal & Zickar, 2012; Echambi & Hess, 2007; Hayes, 2013)。其实, 中心化前后, 回归系数  $c$  的标准误  $se_c$ 、回归系数  $c$  的显著性检验的  $t$  值 ( $t = c / se_c$ ) 也不变, 测定系数  $R_1^2$  也不变, 从而调节效应的效果量  $\Delta R^2$  也不变 (Dalal & Zickar, 2012; Hayes, 2013)。

综上所述, 中心化减少多重共线性对调节效应检验结果没有任何影响, 因此, 中心化减少的共线性也称为非本质的共线性 (nonessential collinearity) (Alike & West, 1991; Dalal & Zickar, 2012; Fairchild & McQuillin, 2010)。真正能影响调节效应检验结果的是自变量  $X$  和调节变量  $Z$  之间的共线性, 被称为本质的共线性 (essential collinearity), 不过, 中心化无

法减少这种本质的共线性 (Alike & West, 1991; Dalal & Zickar, 2012)。

## 2.2 中心化能改善对结果的理解

回归方程 (1) 中, 回归系数  $a$  的含义是, 当调节变量  $Z$  为 0 时, 自变量  $X$  每变化一个单位, 则因变量  $Y$  相应变化  $a$  个单位, 反映的是当调节变量  $Z$  为 0 时, 自变量  $X$  和因变量  $Y$  的关系。如果调节变量  $Z$  的取值范围不包括 0, 即 0 没有实际意义时, 对于回归系数  $a$  的理解就会造成困难。同理, 回归系数  $b$  的理解也存在同样的问题。中心化能有效解决回归系数  $a$  和  $b$  的解释问题, 因为中心化后, 变量的零点移到了原始数据的均值位置。中心化后 (见方程 (3)), 回归系数  $a$  反映的是当调节变量  $Z$  为均值时自变量  $X$  和因变量  $Y$  的关系, 这样, 回归系数  $a$  和  $b$  的统计意义就容易理解了 (Alike & West, 1991; Dalal & Zickar, 2012; Fairchild & McQuillin, 2010; Hayes, 2013)。

## 2.3 小结

在基于多元回归的调节效应分析中, 中心化对调节效应检验结果没有任何影响。但中心化对主效应 (即回归系数  $a$  和  $b$ ) 的检验有影响, 具体表现在两方面, 一是改变回归系数  $a$ 、 $b$  的大小和显著性检验结果 (温忠麟等, 2012; Dalal & Zickar, 2012; Jose, 2012)。二是在自变量  $X$  和调节变量  $Z$  的零点没有实际意义的情况下, 改善对回归系数  $a$  和  $b$  的理解。

# 3 简单斜率检验

## 3.1 选点法

为检验简单斜率  $a+cZ$  是否显著 (即不是 0), 常用如下  $t$  检验:

$$t = \frac{a + cZ}{\sqrt{se_a^2 + 2Z \text{cov}(a, c) + Z^2 se_c^2}} \quad (6)$$

其中,  $se_a$ ,  $se_c$  和  $\text{cov}(a, c)$  分别是方程 (1) 的回归系数  $a$  和  $c$  的标准误和协方差。选择调节变量  $Z$  的某个特定取值 (如  $mean-sd$ ,  $mean$  和  $mean+sd$ ), 即选点, 计算出  $t$  值。然后和  $t_{crit} = t_{.05}(N-K-1)$  进行比较, 其中  $t$  分布的自由度为  $(N-K-1)$ ,  $N$  为样本容量,  $K$  为方程 (1) 的自变量个数 ( $K=3$ )。如果  $t$  绝对值大于  $t_{crit}$ , 就表示在 .05 水平上, 简单斜率  $a+cZ$  显著 (Alike & West, 1991; Cohen et al., 2003)。

为避免繁复的计算, Alike 和 West(1991) 提出了一个替代做法。第一步, 选择调节变量  $Z$  的某个特定取值 (如  $mean+sd$ ), 即选点, 创建一个新

变量  $Z' = Z - (\text{mean} + \text{sd})$ ；第二步，产生乘积项  $XZ'$ ；第三步，进行  $Y = i'' + a''X + b''Z' + c''XZ' + \varepsilon''$  的回归分析，回归系数  $a''$  值就是  $Z = \text{mean} + \text{sd}$  对应的简单斜率值，其显著性检验结果就是所要的简单斜率检验结果。

两种方法都需要首先选点（常选  $\text{mean}$  和  $\text{mean} \pm \text{sd}$ ），故统称为选点法。选点法至少存在以下三点不足。第一，如果调节变量是偏态（skew）分布，则调节变量  $Z$  的特定取值（如  $\text{mean} \pm \text{sd}$ ）可能不在调节变量  $Z$  的取值范围之内，此时的简单斜率检验没有实际意义（Hayes, 2013; Spiller et al., 2013）。第二，选点法得出的结论具有样本特异性（sample specific）。例如两个研究者用调节变量的均值相差较大的两个样本来重复同一个研究，则均值较大样本的  $\text{mean} - \text{sd}$  可能接近于均值较小样本的  $\text{mean} + \text{sd}$ ，最终导致两个研究者的简单斜率检验结果可能不同（Hayes, 2013; Spiller et al., 2013）。第三，选点法选择的是有限的点（常选 2~3 个点），只能了解有限的信息（Bauer & Curran, 2005; Spiller et al., 2013）。为了克服选点法的不足，研究者倡议改用 Johnson-Neyman 法进行简单斜率检验（Bauer & Curran, 2005; Hayes, 2013; Hayes & Matthes, 2009; Jose, 2012; Preacher et al., 2006; Spiller et al., 2013）

### 3.2 Johnson-Neyman 法

Johnson-Neyman 法（简称为 J-N 法）最初用于协方差分析中（Johnson & Neyman, 1936），Bauer 和 Curran (2005) 率先将之拓展到回归模型中，应用于调节变量为连续变量的简单斜率检验。J-N 法的简单斜率检验步骤是，首先确定临界值  $t_{\text{crit}} = t_{.05}(N-K-1)$ 。其次，将公式 (6) 改写为

$$t_{\text{crit}}^2(\text{se}_a^2 + 2Z \text{cov}(a, c) + Z^2 \text{se}_c^2) = (a + cZ)^2 \quad (7)$$

以  $Z$  为主项进行整理得到

$$(t_{\text{crit}}^2 \text{se}_c^2 - c^2)Z^2 + (2t_{\text{crit}}^2 \text{cov}(a, c) - 2ac)Z + (t_{\text{crit}}^2 \text{se}_a^2 - a^2) = 0 \quad (8)$$

记  $A = t_{\text{crit}}^2 \text{se}_c^2 - c^2$ ， $B = 2t_{\text{crit}}^2 \text{cov}(a, c) - 2ac$ ， $C = t_{\text{crit}}^2 \text{se}_a^2 - a^2$ ，则  $Z$  的二元一次方程 (8) 的解为

$$\hat{Z} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \quad (9)$$

$\hat{Z}$  的含义是，当调节变量  $Z$  取值为  $\hat{Z}$  时，简单斜率  $a + c\hat{Z}$  的显著性检验的  $t$  值（由公式 (6) 算出）刚好等于  $t_{\text{crit}}$ ，显著性概率  $p$  刚好等于选定的显著性水平  $\alpha = .05$ ，也就是说， $\hat{Z}$  的简单斜率  $a + c\hat{Z}$  是显著与否的分界点。为了叙述方便，以下将 J-N 法得到的两个解  $\hat{Z}$ （见公式 (9)）称为  $JN_{Z1}$  和  $JN_{Z2}$ （

$JN_{Z1} < JN_{Z2}$ ）。由这两个点就可以得出一个简单斜率  $a + cZ$  显著不为 0 的调节变量  $Z$  的取值区间。值得注意的是，如果  $JN_{Z1}$  和（或） $JN_{Z2}$  值在调节变量的  $[Z_{\min}, Z_{\max}]$  取值范围之外，则  $JN_{Z1}$  和（或） $JN_{Z2}$  将被忽略，以确保调节变量  $Z$  的取值一定在  $Z$  的取值范围内，于是产生表 1 中的 4 种结果（Hayes & Matthes, 2009; Hayes, 2013）：

结果 1：当  $JN_{Z1}$  和  $JN_{Z2}$  都在调节变量  $Z$  的取值范围之内，则可能当调节变量  $Z$  取值区间为  $[JN_{Z1}, JN_{Z2}]$  时，简单斜率显著；也有可能当调节变量  $Z$  取值区间为  $[Z_{\min}, JN_{Z1}]$  和  $[JN_{Z2}, Z_{\max}]$  时，简单斜率显著。

结果 2：只有  $JN_{Z1}$  在调节变量  $Z$  的取值范围之内，则可能当调节变量  $Z$  取值区间为  $[Z_{\min}, JN_{Z1}]$  时，简单斜率显著；也有可能当调节变量  $Z$  取值区间为  $[JN_{Z1}, Z_{\max}]$  时，简单斜率显著。

结果 3：只有  $JN_{Z2}$  在调节变量  $Z$  的取值范围之内，则可能当调节变量  $Z$  取值区间为  $[Z_{\min}, JN_{Z2}]$  时，简单斜率显著；也有可能当调节变量  $Z$  取值区间为  $[JN_{Z2}, Z_{\max}]$  时，简单斜率显著。

结果 4： $JN_{Z1}$  和  $JN_{Z2}$  都不在调节变量  $Z$  的取值范围之内，则可能在调节变量  $Z$  的整个取值区间  $[Z_{\min}, Z_{\max}]$  内，简单斜率都显著，或者反之。

既然每种结果都有两种可能，那么最终结果是哪种情况呢？一个可行的方法是在调节变量  $Z$  的取值范围内选取一些点，利用选点法检验调节变量  $Z$  取这些值时的简单斜率的显著性，利用这些结果就能容易判断最终结果是哪种情况了（Hayes, 2013）。因此，J-N 法实际上包含了选点法，J-N 法的输出结果中包含了选点法的检验结果。

### 3.3 小结

选点法是先选点（即固定调节变量  $Z$  的若干值），然后计算  $t$  值，最后判断简单斜率的显著性。J-N 法则相反，首先固定  $t$  值等于临界值（ $t = t_{\text{crit}}$ ），然后找到简单斜率显著与否的分界点（即对应的  $Z$  值）。选点法一次只能检验一个点（调节变量  $Z$  的某个取值）的简单斜率的显著性，因此 Spiller 等人 (2013) 形象地称之为闪光灯检验（spotlight test）。J-N 法不是立足于某一个点，而是立足于调节变量  $Z$  的整个取值范围  $[Z_{\min}, Z_{\max}]$ ，就像用探照灯巡视调节变量  $Z$  的整个取值范围，来寻找简单斜率显著不为 0 时调节变量  $Z$  的取值区间，这样做不仅摆脱了选点法需要选点的局限，也能获更多信息，因此 Spiller 等人 (2013) 形象地称之为探照灯分析（floodlight analysis）。目前，

表 1 J-N 法的简单斜率  $\alpha+cZ$  检验结果

类型	$JN_{z_1}$ 、 $JN_{z_2}$ 与 $[Z_{\min}, Z_{\max}]$ 关系	简单斜率显著时, $Z$ 可能的取值区间
结果 1	$Z_{\min} < JN_{z_1} < JN_{z_2} < Z_{\max}$	$[JN_{z_1}, JN_{z_2}]$ 、 $[Z_{\min}, JN_{z_1}]$ 和 $[JN_{z_2}, Z_{\max}]$
结果 2	$Z_{\min} < JN_{z_1} < Z_{\max} < JN_{z_2}$	$[Z_{\min}, JN_{z_1}]$ 或 $[JN_{z_1}, Z_{\max}]$
结果 3	$JN_{z_1} < Z_{\min} < JN_{z_2} < Z_{\max}$	$[Z_{\min}, JN_{z_2}]$ 或 $[JN_{z_2}, Z_{\max}]$
结果 4	$JN_{z_1} < Z_{\min} < Z_{\max} < JN_{z_2}$	$[Z_{\min}, Z_{\max}]$ 或空集

J-N 法已可以在 SPSS、SAS 软件上通过 PROCESS 或 MODPROBE 宏 (<http://www.afhayes.com> 下载) 进行窗口操作的简单斜率检验 (Hayes, 2013; Hayes & Matthes, 2009), 也可以在 R 软件上运行 (Preacher et al., 2006)。J-N 法在常用统计软件上的实现, 极大地促进了 J-N 法在实际应用中的使用 (Hayes, 2013)。因此, 建议在调节变量为连续变量时, 使用 J-N 法进行简单斜率检验; 在调节变量为类别变量或研究者对某个调节变量值感兴趣时, 使用选点法进行简单斜率检验。

#### 4 示例

接下来用一个实际例子演示如何进行基于多元回归的调节效应分析。本例要研究的是情绪智力 ( $Z$ ) 在父子依恋 ( $X$ ) 与攻击行为 ( $Y$ ) 之间起调节作用, 变量及其数据 (379 人) 均来自李霓霓等人 (2009) 的研究。由于调节变量 (情绪智力) 为连续变量 (取值范围 [2.53, 4.89], 均值为 3.78, 标准差为 .434), 因此采用 SPSS 19.0 软件的 PROCESS 宏执行未中心化和中心化的两种调节效应分析, 结果见表 2。

表 2 方程 (1) 和方程 (3) 的基于多元回归的调节效应分析结果

模式	自变量	未标准化系数	标准误	$t$ 值	$P$ 值	VIF	$R^2$
方程(1)	截距	1.459	1.095	1.332	.184		
	$X$	.535	.337	1.589	.133	65.299	.139
	$Z$	.543	.290	1.875	.062	18.275	
	$XZ$	-.211	.087	-2.41	.016	105.727	
方程(3)	截距	2.657	.031	85.975	.000		
	$(X - \bar{X})$	-.261	.045	-5.838	.000	1.150	.139
	$(Z - \bar{Z})$	-.148	.073	-2.031	.043	1.151	
	$(X - \bar{X})(Z - \bar{Z})$	-.211	.087	-2.41	.016	1.010	

注: VIF(variance inflation factor) 表示方差膨胀系数。

首先, 在基于多元回归的调节效应分析中 (见表 2), 乘积项  $XZ$  的系数值 (-0.211)、标准误 (.087)、 $t$  值 (-2.41)、 $p$  值 (.016)、 $R^2$  值 (.139) 在两个方程中都不变, 表明对自变量  $X$  和调节变量  $Z$  是否进行中心化, 并不改变调节效应检验的结果, 即情绪智力在父子依恋与攻击行为之间起调节作用。中心化的作用表现在减少了非本质的共线性 (例如乘积项  $XZ$  的 VIF 从 105.727 降低到 1.01, 见表 2), 还改变了自变量  $X$  和调节变量  $Z$  的回归系数的大小和检验结果 (见表 2)。需要说明的是, 表 2 的结果中没有呈现回归系数的标准化解, 因为方程 (1) 得到的乘积项  $XZ$  的标准化回归系数是不正确的, 方程 (3) 得到的乘积项  $XZ$  的标准化回归系数才是正确的, 具体解释请参见相关文献 (温忠麟, 侯杰泰, Marsh, 2008)。

其次, 由于调节效应显著, 因此要进行简单斜率

分析。又因为自变量 (父子依恋) 和调节变量 (情绪智力) 得分的零点没有实际意义, 因此在调节效应分析中采用中心化, 以便于改善对结果的解释。J-N 法的简单斜率检验结果是, 当调节变量  $Z$  在 [3.20, 4.89] 取值时, 简单斜率  $\alpha+cZ=.535-.211Z$  都显著不为 0。

#### 5 讨论与拓展

虽然调节效应分析广泛用于在心理学和其他社科研究领域, 但在基于多元回归的调节效应分析实践中, 仍有部分基本问题亟待澄清。本文首先述评了中心化在基于多元回归的调节效应分析中的作用和简单斜率检验的分析方法, 然后给出了应用建议, 接着用一个例子演示了如何进行调节效应分析。但是, 本文仍然存在一些不足, 尚需进一步深入讨论和拓展。

##### 5.1 均值中心化策略的拓展

第一, 本文只涉及了基于多元回归的调节效应分析中, 自变量和调节变量的中心化, 实际上还有研究者对自变量和调节变量进行标准化。Hayes(2013)用实例证明, 对自变量和调节变量进行中心化和标准化都得到相同的调节效应检验结果, 主要表现在自变量、调节变量和乘积项的回归系数的显著性检验的  $t$  和  $p$  值相同; 对自变量和调节变量进行中心化和标准化都能减少非本质的共线性, 表现在自变量、调节变量和乘积项的 VIF 值相同。对本文示例数据的分析得到和 Hayes(2013) 同样的结果。

第二, 在潜变量的调节效应检验和多层调节效应检验中 (见图 2), 中心化是必要的。因为中心化能通过减少非本质的共线性来减少数据收敛问题, 提高收敛速度 (何晓群, 闵素芹, 2009; Dalal & Zickar, 2012)。中心化在使用某些软件建模时, 还起到简化模型的作用 (吴艳, 温忠麟, 林冠群, 2009; Marsh, Wen, Hau, Little, Bovaird, & Widaman, 2007)。值得注意的是, 在

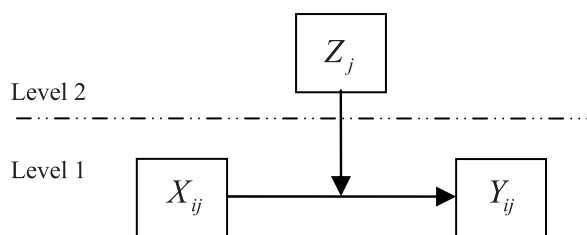


图 2 多层调节效应模型图

以学生 (下标  $i$  表示) 嵌套于班级 (下标  $j$  表示) 的两层调节模型为例,  $\beta_{0j}$  和  $r_{00}$  表示回归方程 (10) 和 (11) 的截距,  $\beta_{1j}$  表示回归方程 (10) 的斜率,  $\varepsilon_{ij}$  和  $\mu_{0j}$  表示回归方程 (10) 和 (11) 的残差, 将方程 (11) 和 (12) 带入方程 (10), 可得方程 (13),

$$Y_{ij} = r_{00} + r_{01}Z_j + \mu_{0j} + r_{10}X_{ij} + r_{11}Z_jX_{ij} + \mu_{1j}X_{ij} + \varepsilon_{ij} \quad (13)$$

如果回归系数  $r_{11}$  显著, 就表示多层调节显著。

将方程 (13) 重新整理得到方程 (14)

$$Y_{ij} = (r_{00} + r_{01}Z_j) + (r_{10} + r_{11}Z_j)X_{ij} + (\mu_{0j} + \mu_{1j}Z_j) + \varepsilon_{ij} \quad (14)$$

其中  $r_{10} + r_{11}Z_j$  就表示简单斜率, 它反映了因变量  $Y_{ij}$  和自变量  $X_{ij}$  的关系是如何受变量  $Z_j$  调节的。J-N 法可以用于简单斜率  $r_{10} + r_{11}Z_j$  的显著性检验, 原理和步骤可见 3.2 部分内容, 只需要将公式 (6)-(9) 中的  $a$  和  $c$  替换成  $r_{10}$  和  $r_{11}$  即可 (Bauer & Curran, 2005; Preacher et al., 2006)。

除了多层调节模型外, Hayes 还将 J-N 法应用到在更复杂的有调节的中介模型和多个调节变量的模型 (Hayes, 2013)、高阶调节变量的模型中 (Hayes & Matthes, 2009)。虽然 J-N 法有如此广阔的应用空

多层调节效应检验中, 中心化又分为总均值中心化 (grand mean-centering) 和组均值中心化 (group mean-centering), 两种中心化策略的作用各有不同, 应根据具体情况选用合适的中心化策略, 例如, 如果研究者关心的是跨层次的调节效应 (见图 2), 建议对层 1 变量进行组均值中心化; 如果研究者只关心层 2 的调节效应, 则建议对层 1 变量进行总均值中心化 (何晓群, 闵素芹, 2009; Enders & Tofghi, 2007)。

## 5.2 简单斜率检验的拓展

第一, 本文只涉及因变量是连续变量的调节效应检验, 实际上简单斜率检验的选点法和 J-N 法还可以推广到因变量是二分类别变量的调节效应检验中, 只需要将公式 (6)-(9) 中的  $t_{crit} = t_{0.05}(N-K-1)$  替换成  $\chi^2_{crit} = \chi^2_{0.05}(1)$  即可 (Hayes & Matthes, 2009)。

第二, 本文只涉及简单调节模型 (如图 1), 实际上还有多层调节模型 (见图 2)。

$$\text{level1} \quad Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij} \quad (10)$$

$$\text{level2} \quad \beta_{0j} = r_{00} + r_{01}Z_j + \mu_{0j} \quad (11)$$

$$\beta_{1j} = r_{10} + r_{11}Z_j + \mu_{1j} \quad (12)$$

间和优势, 但使用 J-N 法进行简单斜率检验在国内才刚刚起步, 仅有侯珂等人 (2014) 在研究团体问题行为 (层 2 调节变量  $Z_j$ ) 对社会网络中心度 (层 1 自变量  $Y_{ij}$ ) 和青少年问题行为 (层 1 因变量  $X_{ij}$ ) 关系的调节作用时, 使用 J-N 法进行简单斜率检验 (侯珂, 邹泓, 刘艳, 金灿灿, 蒋索, 2014)。希望本文能起到抛砖引玉的作用, 相信随着调节研究的深入, 会不断增加我们对调节问题的理解。

## 参考文献

- 何晓群, 闵素芹. (2009). 分层线性模型层 1 自变量中心化问题研究综述. *统计与信息论坛*, 24, 49-52.
- 侯珂, 邹泓, 刘艳, 金灿灿, 蒋索. (2014). 同伴团体对青少年问题行为的影响: 一项基于社会网络分析的研究. *心理发展与教育*, 30, 259-267.
- 李霓霓, 张卫, 李董平, 麦玉娇, 王晶晶, 邢文利. (2009). 青少年的依恋、情绪智力与攻击性行为的关系. *心理发展与教育*, 25, 91-96.
- 吴艳, 温忠麟, 林冠群. (2009). 潜变量交互效应建模: 告别均值结构. *心理学报*, 41, 1252-1259.
- 温忠麟, 侯杰泰, Marsh, W. (2008). 结构方程模型中调节效应的标准化估计. *心理学报*, 40, 729-736.
- 温忠麟, 刘红云, 侯杰泰. (2012). *调节效应和中介效应分析*. 北京: 教育科学出版社.
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting*

- interactions. Newbury Park, CA: Sage.
- Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, 40, 373–400.
- Bohrstedt, G. W., & Goldberger, A. S. (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association*, 64, 1439–1442.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Dalal, D. K., & Zickar, M. J. (2012). Some common myths about centering predictor variables in moderated multiple regression and polynomial regression. *Organizational Research Methods*, 15, 339–362.
- Echambadi, R., & Hess, J. D. (2007). Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science*, 26, 438–445.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12, 121–138.
- Fairchild, A. J., & McQuillan, S. D. (2010). Evaluating mediation and moderation effects in school psychology: A presentation of methods and review of current practice. *Journal of School Psychology*, 48, 53–84.
- Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods*, 41, 924–936.
- Jose, P. E. (2012). *Doing statistical mediation and moderation*. Guilford Press.
- Marsh, H. W., Wen, Z., Hau, K. T., Little, T. D., Bovaird, J. A., & Widaman, K. F. (2007). Unconstrained structural equation models of latent interactions: Contrasting residual- and mean- centered approaches. *Structural Equation Modeling*, 14, 570–580.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31, 437–448.
- Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research*, 50, 277–288.

## Moderation Effect Analyses Based on Multiple Linear Regression

Fang Jie<sup>1</sup>, Wen Zhonglin<sup>2</sup>, Liang Dongmei<sup>2</sup>, Li Nini<sup>3</sup>

(<sup>1</sup> School of Humanities and Communication, Guangdong University of Finance & Economics, Guangzhou, 510320)

(<sup>2</sup> Center for Studies of Psychological Application & School of Psychology, South China Normal University, Guangzhou, 510631)

(<sup>3</sup> Guangzhou Academy of Fine Arts, Guangzhou, 510260)

**Abstract** Moderation indicates that the strength and/or direction of the relation between an independent variable and a dependent variable is affected by a third variable called the moderator. Moderation models are frequently used in the research of psychology and other social science disciplines, still some issues need to be clarified. The purpose of the present study is to clarify two issues in the moderation effect analyses. One is the role of mean-centering for original variables in moderation modeling; the other is the advantages and disadvantages of two existing methods for testing the simple slope.

Firstly, the product term in moderated regression might be collinear with its constituent parts, making it difficult to detect moderation effects. Some researchers presumed that mean-centering could reduce collinearity and improve the precision of estimates from collinear data, but this is not true. After reviewing the role of mean-centering in moderated multiple regression, we emphasize that mean-centering does not change the coefficient of the product term (moderation term) of the regression, but changes the coefficients of the first-order terms (main effect terms) and improves the interpretability of results. Secondly, when a moderation effect is found, the moderation effect needs to be further probed to fully explicate the relationship among the three variables. The most common method for probing the moderation is to test the simple slopes. We discuss the merits and demerits of two methods for testing the simple slopes: the Pick-a-point method and the Johnson-Neyman's method. The Pick-a-point method is to test simple slopes at several specific levels of the predictors and to report whether they are significant or not, whereas the Johnson-Neyman's method is to test simple slopes in the whole range of the predictor and to report the regions in which the simple slopes are significant. We suggest that the Johnson-Neyman's method be adopted to test the simple slopes when the moderator is a continuous variable, and the pick-a-point method be adopted to test the simple slopes when the moderator is a categorical variable or researchers are interested in the test at some special points of the moderator. An example is given to illustrate how to conduct moderation effect analyses by multiple linear regressions and test the simple slopes by using the Johnson-Neyman's method.

Directions for future research on moderation effect analyses are discussed at the end of the paper. In fact, in addition to mean-centering, standardization is an alternative that can be used to analyze moderation effects, and the effect tests with mean-centering and standardization are equivalent. Furthermore, two methods for testing simple slopes can expand to include more complicated moderation models, such as multilevel moderation models and moderation models in which the dependent variable is a binary variable.

**Key words** moderation effect, multiple linear regression, mean-centering, pick-a-point method, Johnson-Neyman's method