

# 1:r 配对设计二值资料一水平多重 Logistic 回归分析

李长平<sup>1,2</sup>, 胡良平<sup>2,3\*</sup>

(1. 天津医科大学公共卫生学院卫生统计学教研室, 天津 300070;

2. 世界中医药学会联合会临床科研统计学专业委员会, 北京 100029;

3. 军事科学院研究生院, 北京 100850

\* 通信作者: 胡良平, E-mail: lphu812@sina.com)

**【摘要】** 本文的目的是介绍“1:1”“1:2”和“1:r”配对设计二值资料一水平多重 Logistic 回归分析方法。统计分析的前提条件是资料值得分析, 而确保配对设计二值资料值得分析的重要基础是“配对设计”科学合理。为此, 本文从六个方面阐述“配对设计的要领”, 不仅介绍了构建“1:1”“1:2”和“1:r”配对设计二值资料一水平多重 Logistic 回归模型的方法, 还通过三个实例, 介绍了采用 SAS 软件实现统计分析结果解释的详细步骤。

**【关键词】** 配对设计; 病例对照; 二值资料; Logistic 回归分析; 优势比

中图分类号: R195.1

文献标识码: A

doi:10.11886/j.issn.1007-3256.2019.04.003

## One-level multiple Logistic regression analysis of the dichotomous choice data collected from the 1:r matched pairs design

Li Changping<sup>1,2</sup>, Hu Liangping<sup>2,3\*</sup>

(1. Department of Health Statistics, School of Public Health, Tianjin Medical University, Tianjin 300070, China;

2. Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing 100029, China;

3. Graduate School, Academy of Military Sciences PLA China, Beijing 100850, China

\* Corresponding author: Hu Liangping, E-mail: lphu812@sina.com)

**【Abstract】** The paper introduced the approach of multiple Logistic regression analysis for binary data collected from the "1:1", "1:2" and "1:r" paired design. The prerequisites of the statistical analysis were that the data were worth analyzing, in order to meet the demand mentioned before, the paired design should be scientific and reasonable. Therefore, this paper emphasized the essentials of paired design from six aspects. This paper not only introduced the approach of constructing the one-level multiple Logistic regression model of the binary data collected from the "1:1", "1:2" and "1:r" paired design, but also introduced the detailed methods of using SAS software package to realize statistical analysis and the result interpretation through three examples.

**【Keywords】** Paired design; Case-control; Binary data; Logistic regression analysis; Odds ratio

## 1 基本概念

### 1.1 引言

在对横断面调查设计二值资料进行统计分析时, 人们都默认一个“隐含假定”是成立的, 即样本中的所有个体都来自“同一个总体”<sup>[1-3]</sup>。然而, 实际可能并非如此。当真实情况偏离那个“隐含假定”较小时, 分析结果的正确性会产生较小的偏离; 如若不然, 则可能得到错误的结果和结论。为此, 人们想到一种简单的办法, 以提高受试对象的“齐同性”, 那就是进行“配对设计”。即从众多的受试对象中找出“条件”最接近的两个或多个, 将它们分到同一个研究组中, 此组就被称为一个配对组(仅含两个个体

或一个配伍组(可含两个以上个体)。此时, 构建的回归模型就叫做“配对设计资料一水平多重 Logistic 回归模型”或“条件一水平多重 Logistic 回归模型”, 简称为“条件 Logistic 回归模型”。

### 1.2 配对设计

#### 1.2.1 配对设计的种类

配对设计有两种, 第一种为“基于定量结果变量”的配对设计, 它是依据基本常识和专业知道确定对研究者关心的“定量结果取值”可能有影响的因素为“配对条件”的一类设计。第二种为“基于二值结果变量”的配对设计, 它是事先确定了结果变量为“二值变量(例如死亡与存活、患病与未患病)”, 再分别从“患病”群体与“未患病(称为对照)”群体中寻找各方面都接近的个体形成“配对组”或

项目基金: 国家高技术研究发展计划课题资助(2015AA020102)

“配伍组”的一类设计。这种设计(指“病例对照配对设计”)很容易与临床统计学中的“病例对照研究设计”相混淆,“病例对照配对设计”是指每个匹配组中都有来自病例组和对照组的个体;而“病例对照研究设计”是指只有两个研究组,一是病例组、二是对照组,显然,它属于一种特殊的“成组设计”。

### 1.2.2 “基于定量结果变量”配对设计的四种形式

“基于定量结果变量”配对设计的四种形式分别为:①自身配对设计(例如,以“个体”服药前后为配对条件,研究两个时间点上测定的疗效指标是否存在统计学差异);②同源配对设计(例如,以“窝别”或“胎别”为配对条件,研究它们在接受两种不同处理后,某定量效应指标上是否存在统计学差异);③条件相近者配对设计(例如,以“可能对定量效应指标有重要影响的一个或多个因素”为配对条件,研究两个独立个体在接受两种不同处理后,某定量效应指标上是否存在统计学差异);④共同生活型配对设计(例如,以“夫妻”为配对条件,研究他们的“智商”是否存在统计学差异)。

### 1.2.3 “基于二值结果变量”配对设计的四种形式

“基于二值结果变量”配对设计的四种形式分别为:①1:1 配对设计(每个配对组中,来自病例组与对照组的个体数各有 1 个);②1:2 配对设计(每个配对组中,来自病例组的个体数有 1 个、来自对照组的个体数有 2 个);③1:r 配对设计(每个配对组中,来自病例组的个体数有 1 个、来自对照组的个体数有 r 个, $r \geq 1$ ,各配对组中的 r 值最好相等,但也可以不等);④m:n 配对设计(每个配对组中,来自病例组的个体数有 m 个、来自对照组的个体数有 n 个, $m \geq 1$ 、 $n \geq 1$ ,各配对组中的 m 值最好相等、n 值最好相等,但也可以不等)。

### 1.2.4 配对设计的要领

配对设计的目的是提高对比组受试对象之间的“齐同性”,从而提高研究的效率和研究结果的正确性。合理的配对设计应把握好以下六个要领:

第一,应注重遗传因素。若所考察的疾病与“遗传”有关,则遗传因素应作为“配对”的首要条件。与患者最具有可比性的“对照”应该从患者的亲属中寻找,若有“双胞胎”则更佳。

第二,应注重环境因素。若所考察的疾病与“环境”有关,则环境因素(如作业环境、生活环境等)应作为“配对”的首要条件。与患者最具有可比性的

“对照”应该从与患者处于相似环境的群体中寻找。

第三,应注重素质因素。若所考察的疾病与“素质”有关,则素质因素(如受教育程度、心理素质、出身背景等)应作为“配对”的首要条件。与患者最具有可比性的“对照”应该从与患者具有相同或相近“素质”的群体中寻找。

第四,应注重行为因素。若所考察的疾病与“行为”有关,则行为因素(如生活方式、生活习惯、风俗民情等)应作为“配对”的首要条件。与患者最具有可比性的“对照”应该从与患者具有相同或相近“行为”的群体中寻找。

第五,应注重精神因素。若所考察的疾病与“精神”有关,则精神因素(如紧张、急躁、抑郁)应作为“配对”的首要条件。与患者最具有可比性的“对照”应该从与患者具有类似“精神状态”的群体中寻找。

第六,应注重心理因素。若所考察的疾病与“心理”有关,则心理因素(如猜疑、嫉妒、心胸狭窄等)应作为“配对”的首要条件。与患者最具有可比性的“对照”应该从与患者具有相同或相近“心理状态”的群体中寻找。

### 1.3 关于本文的题目

本文的题目可能稍显“冗长”,其最精简的表述为“条件 Logistic 回归分析”。之所以采用现在这个题目,是希望把“设计类型(指‘1:r 配对设计’)”“资料类型(指‘二值资料’)”“水平数(指‘一水平’)”“自变量数目(指‘多重’)”和“回归分析类型(指‘Logistic 回归分析’)”都一目了然地呈现出来。

## 2 配对设计二值资料一水平多重 Logistic 回归模型的构建

### 2.1 1:1 配对设计二值资料一水平多重 Logistic 回归模型的构建

假设 A、B 二人是一个配对,他们各自发病的概率分别为  $P_A$  和  $P_B$ ;在有一人发病的情况下,A 发病的概率为  $P$ ,B 发病的概率为  $P'$ 。由于 A、B 中只有一人发病,因此,A 发病和 B 发病是一对互不相容事件;又由于在病例对照研究中 A 发病和 B 发病之间没有联系,因此,它们还是一对独立事件<sup>[4-6]</sup>。设 A 为病例,B 为对照,根据 Bayes 条件概率公式,在只有一人发病的情况下,发病的这个人恰好是病例 A 的条件概率  $P$  可由式(1)计算:

$$P = \frac{P_A(1 - P_B)}{P_A(1 - P_B) + (1 - P_A)P_B} \quad (1)$$

假设  $P$  与是否发病的影响因素  $X$  的关系符合 Logistic 回归模型,则可用式(2)来计算:

$$P = \frac{\left[ \frac{\exp(\beta_0 + \beta x_A)}{1 + \exp(\beta_0 + \beta x_A)} \right] \left[ \frac{1}{1 + \exp(\beta_0 + \beta x_B)} \right]}{\left[ \frac{1}{1 + \exp(\beta_0 + \beta x_B)} \right] + \left[ \frac{1}{1 + \exp(\beta_0 + \beta x_A)} \right] \left[ \frac{\exp(\beta_0 + \beta x_B)}{1 + \exp(\beta_0 + \beta x_B)} \right]} \tag{2}$$

回归模型(2)经过转换后可得:

$$\ln \frac{P}{1-P} = \beta(x_A - x_B) \tag{3}$$

与非配对 Logistic 回归模型相比,截距项( $\beta_0$ )被消去了,这是二者的一个重要区别。

### 2.2 1:r 配对设计二值资料—水平多重 Logistic 回归模型的构建

1:r 配对设计资料的 Logistic 回归分析的原理与 1:1 配对设计基本相同,在模型形式、参数估计以及参数的假设检验方面两者都是一致的。在 1:r 配对设计中,当每个匹配组中的对照数大于 1 时,条件概率及条件似然函数的形式会比 1:1 配对设计略为复杂,下面仅对此进行讨论。

设向量  $X_{it} = (x_{it1}, x_{it2}, \dots, x_{itp})$  表示第  $i$  个匹配组中第  $t$  个观察对象影响因素的观察值。在第  $i$  个匹配组中的  $(1+r)$  个观察对象有 1 名病例的条件下,恰好第一个观察对象属于病例组的条件概率见式(4):

$$P_i = \frac{P(X_{i0}|Y=1) \prod_{i=1}^r P(X_{ii}|Y=0)}{\sum_{i=0}^r [P(X_{ii}|Y=1) \prod_{i=0, i \neq i}^r P(X_{ii}|Y=0)]} \tag{4}$$

它等于观察到的第一组危险因素属于病例而其他危险因素属于对照的概率与各种可能情况下的概率之和的比值。根据 Bayes 原理,式(4)可写成式(5):

$$P_i = \frac{P(Y=1|X_{i0}) \prod_{i=1}^r P(Y=0|X_{ii})}{\sum_{i=0}^r [P(Y=1|X_{ii}) \prod_{i=0, i \neq i}^r P(Y=0|X_{ii})]} \tag{5}$$

将一般 Logistic 回归模型表达式(6)代入式(5),经过化简后得到式(7):

$$P(Y=1|X_{ii}) = P_i(Y=1|X_{ii}) = \frac{1}{1 + \exp[-(\beta_{0i} + \sum_{k=1}^p \beta_k x_{iik})]} \tag{6}$$

$$P_i = \frac{1}{1 + \sum_{i=1}^r \exp[\sum_{k=1}^p \beta_k (x_{iik} - x_{i0k})]} \tag{7}$$

综合  $h$  个匹配组,得到总的条件似然函数表

达式(8):

$$L = \prod_{i=1}^h \frac{1}{1 + \sum_{i=1}^r \exp[\sum_{k=1}^p \beta_k (x_{iik} - x_{i0k})]} \tag{8}$$

有了条件似然函数,就可对其取对数变换,然后对其中的未知参数求偏导数,可获得联立方程组。进而采取合适的非线性迭代算法,便可求得参数的估计值。

【说明】配对设计二值资料多重 Logistic 回归模型的求解非常复杂,一般需要借助统计软件来实现,因篇幅所限,此处从略。

## 3 实例分析

### 3.1 对 1:1 配对设计二值资料的分析

#### 3.1.1 问题与数据

【例 1】为研究发生小于胎龄儿的影响因素,对研究对象的生育史和身高、体重进行匹配,采用 1:1 配对设计,收集 80 对病例和对照的资料<sup>[2-3]</sup>。研究中探讨的影响因素较多,本例仅选取其中 3 个:孕妇年龄( $X_1$ : <25 岁取值为 1, 25 岁 ≤ 年龄 < 30 岁取值为 2, ≥30 岁取值为 3), 孕妇受教育程度( $X_2$ : 小学及以下取值为 1, 中学取值为 2, 大学及以上取值为 3), 家庭人均月收入( $X_3$ : ≤800 元取值为 1, 800 元 < 人均月收入 ≤1500 元取值为 2, >1500 元取值为 3)。见表 1。

表 1 分娩小于胎龄儿影响因素的 1:1 配对病例对照研究资料

配对 编号	年龄	受教育 程度	人均月 收入	配对			
				编号	年龄	受教育 程度	人均月 收入
	病例组(Y=1)			对照组(Y=0)			
1	3	2	2	1	2	2	3
2	2	2	1	2	3	3	2
3	2	2	2	3	1	2	3
4	1	1	1	4	2	2	2
...	...	...	...	...	...	...	...
77	1	1	2	77	2	2	3
78	1	2	1	78	1	1	2
79	2	1	1	79	1	2	2
80	2	2	3	80	2	2	2

注:数据较多,此处从略,详细数据见文献[5]P<sub>187-189</sub>

3.1.2 所需要的 SAS 程序

```
data a;
infile 'D:\sastjfx\data10_1.txt';
input id X1-X3 Y @@;
run;
proc logistic descending;
class X1(ref='3') X2(ref='3') X3(ref='3');
model Y=X1 X2 X3/rsq cl stb selection=stepwise sle=
0.3 sls=0.05;
strata id;
run;
quit;
```

【程序说明】将表 1 中的 80 行 4 列数据(包括配对编号)输入计算机并保存为文本文件格式,取名为“data10\_1.txt”,存储地址为“D:\sastjfx\”(即 D 盘文件夹名为 sastjfx)。过程步语句中的“descending”是希望求“Y=1(患病)”发生概率的回归模型。资料中的三个自变量都是“多值有序自变量”,通常,分析者将它们视为“计量自变量”,直接代入计算。但是,将它们视为“多值名义自变量”进行“哑变量变

换”更合适<sup>[7]</sup>。现在都以它们的第 3 水平为“基准或参照”水平,通过“class 语句”来实现。

“Model 语句”中的选项“rsq”“cl”和“stb”分别要求输出广义决定系数、回归系数的 95% 置信区间和标准化回归系数;最后三个选项是采用逐步回归筛选自变量,其中,“sle=0.3”是选择自变量进入回归模型的显著性水平,取值较大(可称为“宽进”);“sls=0.05”是保留自变量在回归模型中的显著性水平,取值较小(可称为“严出”)。

“strata 语句”要求以“配对编号 id”为分层因素,在本质上就是在每个配对组中求每个变量的“差值”,再利用各“差值”作为“自变量”进行回归建模。

3.1.3 主要分析结果及解释

3 型效应分析

效应	自由度	Wald 卡方	Pr > 卡方
X2	2	13.5636	0.0011

以上是经逐步回归筛选自变量后,仅保留 X<sub>2</sub> 在模型中,因 wald  $\chi^2=13.5636, P=0.0011$ ,说明依据自变量 X<sub>2</sub> 来预测“Y=1”发生的概率是具有统计学意义的。

条件最大似然估计值的分析

参数	自由度	估计值	标准误差	Wald 卡方	Pr>卡方	标准化估计
X2	1	2.7578	0.7586	13.2181	0.0003	0.7860
X2	2	-0.2656	0.4960	0.2868	0.5923	-0.0982

以上结果表明:相对于 X<sub>2</sub> 的第 3 个水平而言,其第 1 水平和第 2 水平的回归系数分别为 2.7578、-0.2656。注意:第 3 个水平的回归系数为“-[2.7578+(-0.2656)]=-2.4922”。-0.2656 与 0 之间差别无统计学意义(P=0.5923),但不宜将其删除。因为一个原变量的全部哑变量应同时保留或同时剔除。其对应的多重 Logistic 回归模型可表示如下:

$$P(Y = 1) = \frac{e^{2.7578X_{21} - 0.2656X_{22}}}{1 + e^{2.7578X_{21} - 0.2656X_{22}}}$$

优比估计值

效应	点估计值	95% Wald 置信限
X2 3-1	190.576	11.177 >999.999
X2 3-2	9.268	1.180 72.793

以上是“X<sub>2</sub>”的两个“哑变量”的优势比“OR”的点估计值及其基于“Wald 算法”的 95% 置信区间。这两个 OR 值是按下面的方法计算出来的:

$$OR_{X_{23-1}} = \frac{e^{2.7578}}{e^{-2.4922}} = e^{2.7578 + 2.4922} = e^{5.25} = 190.566$$

$$OR_{X_{23-2}} = \frac{e^{-0.2656}}{e^{-2.4922}} = e^{-0.2656 + 2.4922} = e^{2.2266} = 9.268$$

【说明】上面输出结果“190.576”与手工计算结果“190.566”的偏差,系计算过程中四舍五入造成的。

参数估计和 Wald 置信区间

参数	估计值	95% 置信限
X2	1	2.7578 1.2711 4.2446
X2	2	-0.2656 -1.2378 0.7065

以上是关于回归系数的基于“Wald 算法”的 95% 置信区间的计算结果。

本例的专业结论:X<sub>2</sub> 为孕妇“受教育程度”,其取值由小到大代表受教育程度由低到高。计算时,以“X<sub>2</sub> = 3(高)”为基础水平,因“OR<sub>X<sub>2</sub>=1</sub> = 190.576”,提示“受教育程度”最低的孕妇生出小于胎龄儿的风险是“受教育程度”最高的孕妇生出小于胎龄儿的风险的 190 倍之多;又因“OR<sub>X<sub>2</sub>=2</sub> = 9.268”,提示“受教育程度”中等的孕妇生出小于胎龄儿的风险是“受教育程度”最高的孕妇生出小于胎龄儿的风险的 9 倍之多。

### 3.2 对 1:2 配对设计二值资料的分析

#### 3.2.1 问题与数据

【例 2】沿用例 1 中的问题和数据,再添加 1 个对

表 2 小于胎龄儿影响因素的 1:2 配对病例对照研究资料

配对 编号	病例组(Y=1)			对照组 I (Y=0)			对照组 II (Y=0)		
	年龄	受教育程度	人均月收入	年龄	受教育程度	人均月收入	年龄	受教育程度	人均月收入
1	3	2	2	2	2	3	1	3	3
2	2	1	1	3	3	2	1	3	3
3	2	2	2	1	2	3	1	3	3
...	...	...	...	...	...	...	...	...	...
18	1	2	2	1	3	1	1	3	2
19	1	3	1	2	2	2	1	2	2
20	1	2	2	1	2	2	1	3	2

注:数据较多,此处从略,详细数据见文献[5]

#### 3.2.2 所需要的 SAS 程序

```
data a;
infile 'D:\sastjfx\data10_2.txt';
input id X1 X2 X3 Y @@;
run;
proc logistic descending;
class X1(ref='3') X2(ref='3') X3(ref='3');
model Y=X1 X2 X3/selection=backward sls=0.05;
strata id;
run;
quit;
```

【程序说明】与例 1 相似,此处从略。

#### 3.2.3 主要分析结果及解释

##### 3 型效应分析

效应	自由度	Wald 卡方	Pr > 卡方
X3	2	8.5076	0.0142

##### 条件最大似然估计值的分析

参数	自由度	估计值	标准误	Wald 卡方	Pr > 卡方
X3 1	1	1.9976	0.7242	7.6079	0.0058
X3 2	1	-0.4759	0.4533	1.1025	0.2937

##### 优比估计值

效应	点估计值	95% Wald 置信限
X3 3-1	33.759	3.166 359.931
X3 3-2	2.845	0.711 11.392

读者可参照例 1 中关于输出结果的解释,尝试解释上面的结果,此处从略。

照组,使其成为 1:2 配对设计,为方便描述,仅选择了 20 个配对组资料,结果仅用于举例说明,有关资料见表 2。

### 3.3 对 1:r 配对设计二值资料的分析

#### 3.3.1 问题与数据

【例 3】为研究视网膜母细胞瘤的危险因素,某医生在两年内共收集病例 8 例。对患者的性别、年龄、民族进行匹配,在同一病区选择了 32 例患有非恶性肿瘤的患者作为对照,对患者母亲孕期吸烟、接受 X 线照射、阴道出血情况(有此类情况赋值为 1,无此类情况赋值为 0)进行调查。见表 3。

#### 3.3.2 所需要的 SAS 程序

所需要的 SAS 程序与例 2 基本相同,此处从略。读者可以参照前面的 SAS 程序,尝试编写 SAS 程序并运行。

#### 3.3.3 主要分析结果及解释

输出内容和格式与例 1、例 2 相同,此处从略。

## 4 讨论与小结

无论是 1:1 配对设计、1:2 配对设计、还是 1:r 配对设计,都应确保同一个匹配组中“对照”所对应的“Y”都取同一种结果(例如,设“病例”为“Y=1”,那么,同一个配对编号所对应的若干个“对照”都应取成“Y=0”)。

应采用“class 语句”将“多值名义变量”和/或“多值有序变量”转换成“哑变量”后,再代入回归模型。基于计算结果,呈现回归模型和解释“优势比 OR”时,应特别慎重;选择“多值名义变量”和/或“多值有序变量”的“基准或参照水平”时,应尽可能确保由某原变量产生的多数哑变量的“OR”估计值的绝对

表 3 视网膜母细胞瘤发病影响因素的 1:4 配对病例对照研究资料

配对 编号	母孕期吸烟	母孕期接受 X 线照射	母孕期阴道出血	配对 编号	母孕期吸烟	母孕期接受 X 线照射	母孕期阴道出血
	病例组 (Y=1)				对照组 I (Y=0)		
1	1	1	0	1	1	0	1
2	1	1	0	2	1	0	0
3	1	1	0	3	1	0	0
4	1	1	0	4	0	0	0
5	1	0	0	5	1	1	0
6	1	1	0	6	1	0	0
7	0	1	0	7	1	0	1
8	1	1	1	8	0	0	0
对照组 II (Y=0)				对照组 III (Y=0)			
1	1	0	0	1	0	0	0
2	1	0	0	2	0	0	0
3	0	0	0	3	0	0	0
4	1	1	0	4	1	0	0
5	1	0	0	5	1	0	0
6	1	0	0	6	0	0	1
7	0	0	0	7	1	0	0
8	1	0	0	8	1	1	0
对照组 IV (Y=0)				对照组 IV (Y=0)			
1	0	0	0	5	0	0	0
2	1	0	0	6	1	0	0
3	0	0	0	7	1	0	0
4	1	0	0	8	1	0	1

注:在 1:r 配对设计中,每个病例组中的“对照数目”最好都是 r,但也可以小于 r

值大于 1,这样便于作出专业解释。选择的思路是:在“class 语句”中,先选择“一水平”为“基准水平”,若结果便于解释,即可;若不便于解释,再选择“最后的一个水平”;若还是不够理想,可尝试选择“中间的某个水平”。

本文从试验设计角度,介绍了提高“配对设计”质量的若干要领以及“1:1”和“1:r”配对设计二值资料多重 Logistic 回归模型的构建原理,并通过三个实例,呈现了“1:1”“1:2”和“1:4”配对设计二值资料的数据结构;基于 SAS 软件,给出了分析结果,并对结果作出了解释。

参考文献

[1] 陈邦定,彭东桃,阳波,等.常德地区严重精神障碍患者暴力

攻击行为研究[J].四川精神卫生,2019,32(1):53-57.

[2] 弋可,曾强,吴俊林,等.绵阳市 COPD 患者焦虑与抑郁检出情况及相关因素分析[J].四川精神卫生,2018,31(6):526-530.

[3] 朱意平,李春阳,陈红红,等.住院精神分裂症患者合并代谢综合征的影响因素[J].四川精神卫生,2018,31(6):540-543.

[4] 王济川,郭志刚. Logistic 回归模型——方法与应用[M].北京:高等教育出版社,2001:1-182.

[5] 胡良平.医学统计学:运用三型理论进行现代回归分析[M].北京:人民军医出版社,2010:186-202.

[6] 胡良平.面向问题的统计学——(2)多因素设计与线性模型分析[M].北京:人民卫生出版社,2012:433-450.

[7] 胡良平.提高回归模型拟合优度的策略(I)——哑变量变换与其他变量变换[J].四川精神卫生,2019;32(1):1-8.

(收稿日期:2019-08-01)

(本文编辑:吴俊林)