

合理进行均值比较——重复测量设计 定量资料方差分析

胡纯严¹, 胡良平^{1,2*}

(1. 军事科学院研究生院, 北京 100850;

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

*通信作者: 胡良平, E-mail: lphu927@163.com)

【摘要】 本文目的是介绍与混合模型有关的5个基本概念、计算方法、一个临床试验实例及其SAS实现。基本概念包括固定效应、随机效应、随机系数、混合模型和多水平模型。计算方法涉及混合模型的经典计算方法和贝叶斯计算方法。临床试验实例涉及三种药物对哮喘患者呼吸功能的影响。本文给出了采用SAS处理实例中定量资料的全过程,包括基于mixed过程和bglimm过程构建6个重复测量设计定量资料混合模型,对输出结果作出了解释,并基于模型拟合效果评价统计量,对所构建的多个混合模型进行了比较,得出了适合本例资料的最优混合模型。

【关键词】 固定效应;随机效应;随机系数;重复测量;混合模型

中图分类号:R195.1

文献标识码:A

doi:10.11886/scjsws20230201002

Reasonably carry out mean value comparison: analysis of variance of the quantitative data of repeated measurement design

Hu Chunyan¹, Hu Liangping^{1,2*}

(1. Graduate School, Academy of Military Sciences PLA China, Beijing 100850, China;

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

*Corresponding author: Hu Liangping, E-mail: lphu927@163.com)

【Abstract】 The purpose of this article was to introduce 5 basic concepts, calculation methods, a clinical trial example and its SAS implementation related to the mixed model. The basic concepts include fixed effects, random effects, random coefficients, mixed models, and multilevel models. The calculation methods involved the classical calculation method and the Bayesian calculation method of the mixed model. An example of the clinical trial involved the effect of three drugs on the ability to breathe in people with asthma. This paper gave the whole process of using SAS software to implement the quantitative data in the example, including constructing 6 repeated measurement design quantitative data mixed models based on the mixed procedure and the bglimm procedure. The output results were explained, and all the mixed models constructed were compared based on the evaluation statistics of the model fitting effect, and the optimal mixed model suitable for the data in the paper was obtained.

【Keywords】 Fixed effect; Random effect; Random coefficient; Repeated measurement; Mixed model

在医学研究中,具有重复测量的定量资料出现的频率很高,通常,重复测量的因素只有一个,但也可以有多个。重复测量因素的数量越多,对其进行方差分析的方法就越复杂。由于重复测量定量资料不同观测点之间往往具有复杂的方差和协方差结构,故需要采用多种不同的混合模型进行处理,进而基于模型拟合统计量对多个混合模型进行比较。本文基于一般混合模型和多水平混合模型,拟合来自重复测量设计的定量资料^[1-2],并对模型的拟合效果进行比较。采用基于贝叶斯理论和马尔科夫链蒙特卡罗方法对拟合优度最好的混合模型进行建模^[3-4],以便给出固定效应、随机效应和随机系数的估计值及其95%最高后验密度(HPD)区间^[1]。

1 基本概念

1.1 固定效应

在常规的试验研究中,试验因素的水平是事先确定的,而不是随机选取的。例如,在一个涉及药物种类的临床试验研究中,研究者拟考查试验药、标准药和安慰剂对某疾病的治疗效果。三种药物在各种条件基本相同的三个足够大的群体中所产生的疗效是相对确定的,在统计学上就称“药物种类”是一个“固定效应因素”。

1.2 随机效应

假设有若干种治疗某疾病的方法,考虑到人

力、物力、财力和时间问题,研究者拟从众多治疗方法中随机选取三种进行临床试验,将符合临床要求的全部受试者随机均分为三组(每组样本含量相等且数量符合统计学要求),分别采用三种方法进行“治疗”。此时,“治疗方法”这个因素的3个水平中的每个水平都有1个固定的效应值,还有1个反映众多治疗方法之间疗效变异性大小的数量,称为随机效应方差。因此,被随机选取的三种治疗方法就称为“随机效应因素”。

1.3 随机系数

在一个由南方成年人和北方成年人组成的总体中,研究体重Y随身高X变化的依赖关系时,最简单的统计分析方法是采用直线回归分析。此时,假定得到的直线回归方程见式(1)。也可以分别在南方成年人与北方成年人各自的总体中进行随机抽样,仍然采用直线回归方程来研究Y对X的依赖关系,可以得到两个直线回归方程,见式(2)、式(3)。

$$\hat{Y} = a_{\text{总}} + b_{\text{总}}X \quad (1)$$

$$\hat{Y} = a_{\text{南方}} + b_{\text{南方}}X \quad (2)$$

$$\hat{Y} = a_{\text{北方}} + b_{\text{北方}}X \quad (3)$$

由基本常识可知,以上3个式子中的3个截距可能不相等,3个斜率也可能不相等。如果仅有截距不等,而斜率之间的差异无统计学意义,在统计学上就用一个截距带随机效应的直线回归方程来描述整个总体中Y关于X的直线回归方程,并称其为“随机截距直线回归方程”;同理,当上述3个公式中的截距和斜率之间差异均有统计学意义时,在统计学上就需要用一个截距带随机效应和一个斜率带随机效应的直线回归方程来描述整个总体中Y关于X的直线回归方程,并称其为“随机截距-随机斜率直线回归方程”。由于截距和斜率都是直线回归方程中的回归系数,故此时的回归方程又被称为“随机系数直线回归方程”^[5]。

1.4 混合模型

若一个回归模型中既包含固定效应变量或因素,又包含随机效应变量或因素,在统计学上就将这样的回归模型称为混合效应模型,简称为混合模型。

1.5 多水平模型

若一个总体可以按照某个或某些因素被分层,但仍希望采用一个统一的回归模型来描述因变量随自变量变化的依赖关系,此时,就需要构建随机

系数回归模型。这种回归模型也被称为多水平回归模型,简称为多水平模型。

2 计算方法

2.1 混合模型的经典计算方法

2.1.1 混合模型的公式

一般线性模型是通过glm过程拟合的模型。然而,很多情况下,关于误差 ε 的分布假设过于限制。混合模型通过允许更灵活地指定误差 ε 的协方差矩阵来扩展一般线性模型。混合模型表达式见式(4)。

$$y = X\beta + Z\gamma + \varepsilon \quad (4)$$

式(4)中,除了添加已知设计矩阵Z和未知随机效应参数向量 γ 之外,一切都与一般线性模型中的相同。矩阵Z可以包含连续变量或虚拟变量。混合模型的名称来源于这样一个事实,即该模型同时包含固定效应参数 β 和随机效应参数 γ 。

上述分析中的一个关键假设是 γ 和 ε 服从正态分布, γ 和 ε 的期望和方差分别见式(5)、式(6)。

$$E \begin{bmatrix} \gamma \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (5)$$

$$\text{Var} \begin{bmatrix} \gamma \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \quad (6)$$

因此, y 的方差为 $V=ZGZ'+R$ 。用户可以通过设置随机效应设计矩阵Z并指定G和R的协方差结构来对V建模。

2.1.2 混合模型中协方差参数的估计

混合模型的参数估计比一般线性模型的参数估计更困难。不仅有参数 β ,而且在 γ 、G和R中也有未知参数。最小二乘法不再是最佳方法。广义最小二乘法(GLS)更合适,其对应的表达式见式(7)。

$$(y - X\beta)' V^{-1} (y - X\beta) \quad (7)$$

在式(7)中,由于缺乏关于V、G、R的信息,无法给出式(7)的估算结果。一种方法是使用估计的GLS,用户可以将V的一些合理估计插入最小化问题中。因此,目标是找到G和R的合理估计值。在许多情况下,统计学家利用 γ 和 ε 是正态分布的假设^[6],使用基于似然的方法来实现对式(7)的估计。Proc mixed过程实现了两种基于似然的方法:即最大似然法(ML)和限制性最大似然法(REML)。这两种似然法的一个有利的理论性质是它们允许容纳随机缺失数据^[7-8]。

Proc mixed过程构造与ML或REML有关的目的

标函数,并在所有未知参数上使目标函数最大化。使用微积分,可以将这个最大化问题简化为仅覆盖 G 和 R 中的参数。相应的对数似然函数分别见式(8)、式(9)。

$$ML: l(G, R) = -\frac{1}{2} \log|V| - \frac{1}{2} r'V^{-1}r - \frac{n}{2} \log(2\pi) \quad (8)$$

$$REML: l_r(G, R) = -\frac{1}{2} \log|V| - \frac{1}{2} \log|X'V^{-1}X| - \frac{1}{2} r'V^{-1}r - \frac{n-p}{2} \log(2\pi) \quad (9)$$

上式中, $r = y - X(X'V^{-1}X)^{-1}X'V^{-1}y$, p 是 X 的秩。通过使用 Newton-Raphson 算法, Proc mixed 过程可使上述两式达到最小化。

2.1.3 混合模型中固定效应和随机效应的估计

ML、REML、MIVQUE0 或 Type1、Type3 提供 G 和 R 的估计值,分别表示为 \hat{G} 和 \hat{R} 。为了获得 β 和 γ 的估计值,标准方法是求解混合模型方程,见式(10)。

$$\begin{bmatrix} X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\ Z'\hat{R}^{-1}X & Z'\hat{R}^{-1}Z + \hat{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} X'\hat{R}^{-1}y \\ Z'\hat{R}^{-1}y \end{bmatrix} \quad (10)$$

解决方案也可以写成式(11)和式(12)的形式。

$$\hat{\beta} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y \quad (11)$$

$$\hat{\gamma} = \hat{G}Z'\hat{V}^{-1}(y - X\hat{\beta}) \quad (12)$$

如果 \hat{G} 是奇异的,则混合模型方程可修改如下,见式(13)。

$$\begin{bmatrix} X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z\hat{G} \\ \hat{G}'Z\hat{R}^{-1}X & \hat{G}'Z\hat{R}^{-1}Z\hat{G} + G \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} X'\hat{R}^{-1}y \\ \hat{G}'Z\hat{R}^{-1}y \end{bmatrix} \quad (13)$$

分别用 C 和 M 表示混合模型方程的非奇异 \hat{G} 和奇异 \hat{G} 形式的广义逆。在非奇异情况下,解 $\hat{\gamma}$ 直接估计随机效应;但在奇异情况下,通过反向变换 $\hat{\gamma} = \hat{G}\hat{\tau}$ 实现对随机效应的估计,其中 $\hat{\tau}$ 是修正混合模型方程的解。类似地,虽然在非奇异情况下 C 本身是 $(\hat{\beta}, \hat{\gamma})$ 的估计协方差矩阵,但在奇异情况下, $(\hat{\beta}, \hat{G}\hat{\tau})$ 的协方差估计由 PMP(说明:这是 P、M、P 三个矩阵连乘)给出,其中, M 是混合模型方程式(13)的奇异 \hat{G} 形式的广义逆;而 P 由下式给出,见式(14)。

$$P = \begin{bmatrix} I & \\ & \hat{G} \end{bmatrix} \quad (14)$$

2.2 混合模型的贝叶斯计算方法

Proc bglimm 过程是一种基于模拟的过程,它使用多种采样算法从广义线性混合模型(GLMM)的参数联合后验分布中提取样本。采样方法包括共轭

采样器、直接采样器、Gamerman 算法和无 U 形转弯采样器(No-U-Turn Sampler, NUTS)。

对于已知采样分布中的缺失值采用直接抽样法进行计算;对于非正态模型中的固定效应参数和随机效应参数,需要采用 Gamerman 抽样算法进行计算。对于协方差参数,当共轭抽样算法不可用时,可以采用 NUTS 算法。

Proc bglimm 过程通过吉布斯采样有条件地更新参数。固定效应参数 β 在每次迭代中共同更新。随机效应参数由集群更新,除非未指定 SUBJECT= 选项。在这种情况下,来自同一随机语句的随机效应参数将被联合更新。缺失的数据值按顺序更新, G 侧和 R 侧协方差参数在完全后验条件下分别更新。

混合效应模型的贝叶斯计算方法主要有共轭抽样算法、Gamerman 算法和哈密顿蒙特卡罗(HMC)算法等,这些算法都比较复杂,感兴趣的读者可参阅文献[1]。

3 实例与 SAS 实现

3.1 问题与数据结构

3.1.1 一个临床试验问题及数据

【例 1】该实例来自文献[1],数据由一家制药公司收集,研究三种药物对哮喘患者呼吸功能的影响。治疗药物包括标准药物(a)、试验药物(c)和安慰剂(p)。患者在不同的时间按相同顺序依次接受三种药物治疗。记录患者基线用力呼气量(Fev),并在治疗后的 8 小时内,每小时测量患者的 Fev 值。试验安排和结果见表 1,试分析三种药物对 Fev 的影响差异是否有统计学意义。

表 1 三种药物对患者用力呼气量影响的试验结果
Table 1 Experimental results of the effects of three drugs on forced expiratory volume of patients

患者编号	基线 Fev 值	药物种类	观测时间	Fev
201	2.46	a	1	2.68
⋮	⋮	⋮	⋮	⋮
232	2.88	p	6	2.89

注:同一个患者的编号最多可能会出现 24 次,因为每位患者要先后服用三种药物,服每种药物后要在 8 个时间点上重复观测 Fev 值;具有缺失值的患者的编号出现的次数小于 24

3.1.2 对数据结构的分析

本试验的数据结构:每例患者对应一个“编号(变量名为 Patient)”;从每例患者身上测得 27 个

“Fev 值”,其中,第一组 9 个 Fev 值的变量名分别为“b1、a1、…、a8”,第二组 9 个 Fev 值分别对应于变量名为“b2、c1、…、c8”,第三组 9 个 Fev 值分别对应于变量名为“b3、p1、…、p8”,其中,b1、b2、b3 为三次基线 Fev 值;a、c、p 分别为三种药物,其后数字 1~8 分别代表用药后测量 Fev 值的时间点。由此可知,这个临床试验对应的试验设计类型为“具有一个协变量且具有两个重复测量因素(即药物种类与观测时间)的重复测量设计”。由于各时间点上的观测结果 Fev 值都是定量的,故本例的数据结构为具有前述设计类型下的“定量资料”。值得注意的是,在本例的原始资料中,部分患者在某些时间点上的数据缺失。

3.1.3 创建 SAS 数据集

设所需要的 SAS 数据步程序如下:

```
data Fev;
input Patient Baseval Drug$ Hour Fev;
datalines;
201 2.46 a 1 2.68
201 2.46 a 2 2.76
201 2.46 a 3 2.50
232 2.88 p 5 2.69
232 2.88 p 6 2.89
;
```

【变量说明】Patient 为患者编号,BaseVal 为 Fev 的基线值,Drug 为药物类型(其水平分别为 a、c、p),Hour 为重复观测时间(第 1、2、…、8 小时),Fev 是定量的结果变量。

3.2 用 SAS 实现统计分析

3.2.1 拟构建的 6 个混合效应模型

模型 1:将“Hour”视为“随机效应因素”;模型 2:将“Hour”视为“重复测量因素”;模型 3:将“Hour (Drug)”视为“随机效应因素”;模型 4:将“Hour (Drug)”视为“重复测量因素”;模型 5:将“截距”视为“随机系数”;模型 6:将“截距”视为“随机系数”,将“Hour”视为“随机效应因素”。

3.2.2 基于 mixed 过程构建 6 个重复测量设计定量资料混合模型

3.2.2.1 分析本例资料的 6 个混合效应模型的 SAS 程序

拟合模型 1 所需要的 SAS 程序:

```
proc mixed data=Fev method=ml;
class Drug Hour;
model Fev=BaseVal Drug/s;
random Hour;
run;
```

拟合模型 2 所需要的 SAS 程序:

将模型 1 中 SAS 程序的第 4 句修改为“repeated Hour;”。

拟合模型 3 所需要的 SAS 程序:

将模型 1 中 SAS 程序的第 4 句修改为“random Hour(Drug);”。

拟合模型 4 所需要的 SAS 程序:

将模型 1 中 SAS 程序的第 4 句修改为“repeated Hour(Drug);”。

拟合模型 5 所需要的 SAS 程序:

```
proc mixed data=Fev method=ml;
class Patient Drug Hour;
model Fev=BaseVal Drug Hour;
random int /sub=Patient;
random Hour/ subject=Patient(Drug) type=un;
run;
```

拟合模型 6 所需要的 SAS 程序:

将模型 5 中 SAS 程序的第 5 句修改为“repeated hour / subject=Patient(Drug) type=un;”。

3.2.2.2 上述 6 个混合模型的主要输出结果

上述 6 个混合模型的输出结果包括模型的基本信息、拟合效果评价信息、随机效应和随机截距参数的估计信息以及假设检验信息,因篇幅所限,下面列出有关“拟合效果评价信息”。见表 2。基于“-2logL 值”“固定效应参数个数”“随机效应参数个数”等信息,采取似然比检验法可知,模型 6 对本例资料的拟合效果最好。

3.2.2.3 模型 6 的主要输出结果

【SAS 输出结果及解释】协方差参数估计的计算结果见表 3。结果为随机截距方差的估计值(即 $\sigma_{\mu_0}^2=0.180$)和 8 个时间点两两之间的协方差参数的估计值(共 36 个),为节省篇幅,中间省略了 30 行。

采用模型 6 拟合本例资料的拟合效果评价统计量见表 4。-2logL 值为 52.700,当模型中需估计的参数个数相同时,-2logL 的值越小越好。

对模型中固定效应进行 3 型检验的计算结果见表 5。结果表明,模型中的 3 个固定效应均有统计学意义。

表2 模型拟合过程和拟合效果评价信息摘录

Table 2 Extract of the models fitting process and fitting effect evaluation information

模型编号	模型类别	固定效应	随机效应	语句	固定效应参数个数	随机效应参数个数	-2logL
1	随机效应	BaseVal Drug	Hour	Random	2	2	655.4
2	随机效应	BaseVal Drug	Hour	Repeated	1	4	682.2
3	随机效应	BaseVal Drug	Hour(Drug)	Random	2	4	662.2
4	随机效应	BaseVal Drug	Hour(Drug)	Repeated	1	4	682.2
5	随机效应及截距	BaseVal Drug Hour	Hour Int	###	38	3	52.7
6	随机效应及截距	BaseVal Drug Hour	Hour Int	***	37	3	52.7

注:求解模型中参数估计值的方法为最大似然法(ML),拟合效果均优于采用限制性最大似然法(REML);“###”代表采用了两个Random语句,并使用了“subject=Patient(Drug) type=un”选项;“***”代表采用了一个Random语句和一个Repeated语句,并使用了“subject=Patient(Drug) type=un”选项

表3 协方差参数估计的计算结果

Table 3 Calculation results of covariance parameter estimation

协方差参数	对象	估计
Intercept	Patient	0.180
UN(1,1)	Patient(Drug)	0.099
UN(2,1)	Patient(Drug)	0.054
UN(2,2)	Patient(Drug)	0.066
...
UN(8,6)	Patient(Drug)	0.127
UN(8,7)	Patient(Drug)	0.156
UN(8,8)	Patient(Drug)	0.181

表4 偏差信息准则的计算结果

Table 4 Calculation results of deviation information criteria

统计量	估计值
-2logL	52.700
AIC(赤池信息准则)	148.700
AICC(校正的赤池信息准则)	161.500
BIC(萨瓦-贝叶斯准则)	205.200

表5 固定效应的3型检验的计算结果

Table 5 Calculation results of type 3 test for fixed effects

效应	分子自由度	分母自由度	F	Pr>F
Baseval	1	381	27.560	<0.010
Drug	2	381	59.970	<0.010
Hour	7	381	10.530	<0.010

3.2.3 基于bglimm过程构建前述第6个混合模型

SAS/STAT中的bglimm过程基于贝叶斯理论,并采用马尔科夫链蒙特卡罗算法估计模型中的固定效应参数和随机效应参数的估计值。所需要的SAS程序如下:

```
proc bglimm data=Fev nmc=10000 seed=44672057
outpost=Fev_mod4;
class Drug Patient Hour;
model FEV = BaseVal Drug Hour;
random int / subject=Patient;
```

```
repeated hour/subject=Patient(drug) type=un;
run;
```

【SAS输出结果及解释】模型中各参数的后验汇总及区间的计算结果见表6。输出结果为模型中固定效应参数和随机效应参数的估计值和95%HPD。因篇幅所限,此处省略了很多行数据,还删掉了95%HPD包含0的行(因为这些行上的参数与0之间的差异无统计学意义)。

由表6可知,模型的截距为0.995,95%HPD为(0.166~1.883);基线的回归系数为0.601,95%HPD为(0.302~0.919);药物a和药物c的疗效与安慰剂(p)相比,差异均有统计学意义,其回归系数分别为0.330、0.541,95%HPD分别为(0.172~0.494)、(0.378~0.703);与第8个时间点相比,前4个时间点均有统计学意义;8个时间点上的方差均有统计学意义,任何两个不同时间点之间的协方差无统计学意义;模型误差的方差 $\sigma_e^2=0.347$,其95%HPD为(0.174~0.563)。

表6 后验汇总和区间的计算结果

Table 6 Calculation results of the posterior summaries and intervals

参数	数目	均值	标准差	95%HPD	
Intercept	10 000	0.995	0.442	0.166	1.883
Baseval	10 000	0.601	0.159	0.302	0.919
Drug a	10 000	0.330	0.081	0.172	0.494
Drug c	10 000	0.541	0.083	0.378	0.703
Drug p	0	-	-	-	-
...
Residual UN(7,7)	10 000	0.350	0.069	0.226	0.489
Residual UN(8,8)	10 000	0.346	0.073	0.218	0.489
Random Var	10 000	0.347	0.109	0.174	0.563

3.2.4 结论

基于多种混合模型得到的结论基本相同,即截距、基线值、a药与p药之差量、c药与p药之差量、前4个时间点与第8个时间点之差量均有统计学意义,最合适的统计模型为模型6。

4 讨论与小结

4.1 讨论

将重复测量设计定量资料视为独立重复试验资料,采用一般线性模型分析,属于一种近似做法,所得结果和结论仅供参考。若仅采用具有重复测量设计定量资料的一水平统计模型来分析本例资料(因篇幅所限,有关的SAS程序和详细的输出结果从略),其拟合效果也较差($-2\log L=671.5$, 1个协方差参数和3个固定效应参数, $AIC=673.5$)。拟合本例资料最合适的统计模型为随机截距多重线性混合效应模型,即二水平多重线性混合效应模型(共估计了37个协方差参数,3个固定效应参数, $-2\log L=100.5$, $AIC=174.5$),该模型既考虑了重复测量设计因素,又对资料进行了分层分析。

4.2 小结

本文分别采用一般线性模型、一般线性混合模型和多水平线性混合模型分析了临床资料,还基于贝叶斯理论和马尔科夫链蒙特卡罗方法构建了多水平线性混合模型。得到的结论基本一致,但从模型对资料的拟合效果来看,采用多水平线性混合模型分析本文的临床资料最合适。

参考文献

- [1] SAS Institute Inc. SAS/STAT®15.1 user's guide[M]. Cary, NC: SAS Institute Inc, 2018: 1205-1306, 6533-6728.
- [2] Wang J, Xie H, Fisher JH. Multilevel models: application using SAS[M]. Beijing: Higher Education Press, 2009: 1-112.
- [3] 刘金山, 夏强. 基于MCMC算法的贝叶斯统计方法[M]. 北京: 科学出版社, 2016: 4-117.
Liu JS, Xia Q. Bayesian statistical method based on MCMC algorithm[M]. Beijing: Science Press, 2016: 4-117.
- [4] 康崇禄. 蒙特卡罗方法理论和应用[M]. 北京: 科学出版社, 2015: 86-149.
Kang CL. Monte Carlo method theory and application [M]. Beijing: Science Press, 2015: 86-149.
- [5] Littell RC, Milliken GA, Stroup WW, et al. SAS system for mixed models [M]. Cary, NC: SAS Institute Inc, 1996: 31-170, 253-302.
- [6] Hartley HO, Rao JN. Maximum-likelihood estimation for the mixed analysis of variance model[J]. Biometrika, 1967, 54(1): 93-108.
- [7] Rubin DB. Inference and missing data[J]. Biometrika, 1976, 63(3): 581-592.
- [8] Little RJA. Modelling the drop-out mechanism in repeated measures studies[J]. J Am Stat Assoc, 1995, 90(431): 1112-1121.

(收稿日期:2023-02-01)

(本文编辑:陈霞)