

《心理学报》审稿意见与作者回应

题目：LGM 模型中缺失数据处理方法的比较：ML 方法与 Diggle-Kenward 选择模型

作者：张杉杉 陈楠 刘红云

第一轮

审稿人 1 意见：该文在潜变量增长模型（LGM）中同时考虑数据缺失模式和分布形态对缺失数据处理的两种主要方法（Diggle-Kenward 选择模型和 ML 方法）的影响，这是一项艰苦的研究，因为生成 NMAR 缺失数据和非正态数据都是不容易的事情。研究问题有新意，也有一定的实际意义；研究方法得当，结果合理；文章表述清楚。但文章有改进的空间，问题如下：

意见 1：摘要（也见文中相关内容）：“在 MAR 缺失机制下，基于 NMAR 的 Diggle-Kenward 选择模型仍具有一定的稳健性。”这个算不算稳健性呢？如果正态条件下推出的方法在非正态情形表现也不错，就说方法对（非正态）分布有稳健性。如果在独立的条件下推出的方法在相关的情形表现也不错，就说方法对相关性的稳健性。注意一下，正态是分布的特殊情形，独立是相关的特殊情形，在特殊情形推出的方法，在一般情形也表现不错，说明方法有稳健性。如果是 MAR（随机缺失）推出的方法，在 NMAR 情形也表现不错，才是平常意义上的稳健性，因为 MAR 是 NMAR 的特例。

回应：谢谢审稿专家的意见，已经将“在 MAR 缺失机制下，基于 NMAR 的 Diggle-Kenward 选择模型仍具有一定的稳健性。”修改为“缺失机制对基于 MAR 的 ML 方法有较大的影响。在 MNAR 缺失机制下，基于 MAR 的 ML 方法对 LGM 模型中截距均值和斜率均值的估计不具有稳健性。”

意见 2：公式中（1）中的 i 表示什么要有说明，建议换一个符号，因为 LGM 默认 i 和 s 为截距和斜率。

回应：已经按照审稿专家的建议，作了修改，具体见 3.2 部分的修改。

意见 3：4.1 节，“研究共包括 $4 \times 4 \times 4 \times 2 = 128$ 种模拟条件”，“模拟条件”改成“实验处理”（即实验水平组合），会比较容易明白。

回应：已经按照审稿专家的建议，作了修改，具体见 P9 模拟设计中的修改。

意见 4：4.4 节，“使用 Mplus 软件的 Monte Carlo 模块完成两种模型下的数据分析”。如果没有记错，MPLUS 不像 LISREL 那样可以通过 RP 指令依次读外来数据文件的不同样本。问题是，R 产生的数据（每种处理重复 500 次，相当于 500 个样本），是如何读入 Mplus 软件分析的？文中请给出简要说明。还有，建议写出在 Mplus 中进行两种分析的语句，如果是一句话，可以在正文中说明，如果比较长，在附录中放一个完整的分析程序。

回应：Mplus 的模拟研究中可以直接读取外部的数据文件，已经在文章中进行了说明，具体见 P。同时，在附件中给出了 ML 方法和 DK 方法的语句。

意见 5：6.2 节，“由于 NMAR 缺失机制的理论假设较 MCAR 和 MAR 而言更为严格”，到底谁更严格呢？应当是 MCAR 最严格吧？

回应：已修改为：“由于 NMAR 缺失机制的理论假设较 MCAR 和 MAR 而言更为复杂”。

意见 6：表 1 续表中的表头出现差错。表中数字保留 3 位小数足以满足需要，且方便阅读比较。

回应：谢谢，已修改。

意见 7：文献格式需要按 APA 格式修改。

回应：谢谢，已修改。

意见 8：文字表述方面可以推敲。例如，“如果模型的假设条件不满足（即对于非正态的数据）”，建议改成“如果模型的正态性假设不满足”；“同时关注于其方差”，建议改成“同时关注其方差”。

回应：谢谢，已仔细检查文字表述，修改了部分不清楚的表述。

审稿人 2 意见：追踪研究方法作为探讨心理与行为发生、发展和变化的有效途径之一，已被广泛应用于心理研究的各个领域。本文对 LGM 中缺失数据的处理方法作出介绍，并通过蒙特卡洛模拟研究对 Diggle-Kenward 选择模型和 ML 方法进行比较，得出了相应的结论，是

一篇有一定意义的基础研究论文。但是，文章也存在一些问题：

意见 1：潜变量增长模型中缺失数据的处理方法有许多，为何作者只选取 Diggle-Kenward 选择模型和 ML 方法进行比较？选取的原则或理由，没有交代清楚。

回应：回复：感谢审稿专家的意见，的确如您所言，潜变量增长模型中缺失数据的处理方法有很多，也已经有很多研究就不同的方法进行过比较（Enders, 2010; Jellic, Phelps & Lerner, 2009）。本研究选取 Diggle-Kenward 选择模型和 ML 方法进行比较，主要是基于以下考虑：本研究是在以往研究结论的基础上展开的研究，重点关注追踪研究中随机缺失或非随机缺失情境下，LGM 模型中随机缺失数据的处理的问题。对于随机缺失的数据，已有研究推荐使用 ML 的方法（Yuan, Yang-Wallentin, & Bentler, 2012）；而对于非随机缺失的数据，往往需要基于 LGM 模型假设数据的缺失机制，本研究所考虑的情况是基于选择模型的非随机缺失机制，因此选择了常用的 Diggle-Kenward 选择模型。

关于这一点，修改稿中进一步细化了和部分的内容，详见修改稿引言部分（第四段和第五段的内容）；同时在讨论部分进一步指出未来研究中可以进一步探讨的问题（讨论中建议部分的第（6）条）。

意见 2：缺失数据处理方法的选择依赖于缺失数据产生的机制，缺失机制包括完全随机缺失、随机缺失和非随机缺失，为何作者在模拟研究中只选择随机缺失和非随机缺失两种机制进行方法的比较？

回应：谢谢审稿专家的提问。本研究之所以只考虑了 MAR 和 MNAR 两种缺失机制，主要是基于以下考虑：（1）基于以往大量的研究，对于完全随机缺失（MCAR）的数据大多数的处理方法都能得到较好的结果，ML 方法也能得到精确的估计结果，因此对于 MCAR 的数据缺失方法的选择基本上不存在争议；（2）追踪研究中，数据往往不满足 MCAR 的假设，就目前这一领域的研究趋势来看，研究者面临的问题是如何处理追踪研究中 MNAR 的数据，因此为了研究目的的聚焦，本研究只考虑 MAR 和 MNAR 的情景。

意见 3：如何研究都有其应用意义，本文在这方面未能予以阐述清楚。

回应：谢谢审稿专家的意见。本研究主要探讨了：（1）MAR 和 MNAR 缺失机制下，模型选择的问题；（2）非正态数据分布下，模型参数估计稳健性的问题；（3）考虑数据偏离正态的程度、数据缺失比例、样本量对不同处理方法的影响。这些研究结论可以为实际应用中，研

究者对方法的选择提供一定的依据。基于本研究的结果，研究中提出了实际应用中应该考虑的问题和分析的一般步骤。具体见讨论的建议部分。

意见 4: 其他还有一些写作或专业术语的错漏，如：非随机缺失的英文缩写应为 MNAR (Missing Not at Random, MNAR)，而非 NMAR。

回应: 感谢审稿专家仔细的审阅，已经检查并修改了表述的错误。

第二轮

审稿人 1 意见: 修改稿还有如下问题：

意见 1: 公式 (1) 对分布的描述，通常是不需要加上个体的下标“k”。加上后，很难区分 Y_k 和后面的 Y_t 。要么不要使用 k 下标，要么坚持使用 k 下标，在引入时间下标 t 后，变成双下标。

回应: 谢谢审稿专家的建议，已经按照建议修改，去掉了下表 k。具体修改见 P6，公式 (1) 和 (2) 以及相应的文字解释。

意见 2: 修改稿中说到模拟：“在每一种模拟条件下，运用 Mplus 中的 Monte Carlo 模块对 500 个样本进行分析，在此过程中，只需提供 500 个数据集的文件名列表即可实现批量处理。”在附录中，使用的语句是“FILE=mnarmar40201000replist.dat”。而其中 500 个数据文件来自本 R 语言生成。但附录没有相应的 R 语言程序，希望补充，至少供审稿人参考，不然整个数据生成过程就像一个黑箱（通常的模拟研究，数据生成是靠所用的 SEM 软件，所以不存在这个问题），无从判断模拟的合理性。

回应: 由于生成随机数据的 R 语句较长，作为附件提供，具体见附件中的 R 语句。请审稿专家审阅。

意见 3: 目前的标题过大，其实只是比较了两种主要的缺失值处理方法。

回应: 同意审稿专家的意见，原文的标题确实过大。已经将文章题目修改为：“LGM 模型中缺失数据处理方法的比较及稳健性：ML 方法与 Diggle-Kenward 选择模型”。

审稿人 2 意见：论文研究了潜变量增长中缺失数据处理方法的问题，研究问题比较深入，对理论和实践均有一定的指导意义。表现出中国该领域研究不断进步。但论文存在的问题也比较突出。

意见 1：ML 方法适合于 MAR 情况，而 D-K 方法则是为 MNAR 情况所设计的，直接比较这两种方法，在研究设计上存在逻辑缺陷的，这也是审稿人一所提出的问题，但作者并没能很好的回应该问题；我也建议，如果作者的标题为 LGM 模型中缺失数据处理方法的比较，应加入更多的 MNAR 方法间比较。

回应：谢谢审稿专家的意见。结合审稿专家一的建议，目前文章的标题的确过于宽泛。已经将文章的标题作了修改，修改后的标题为：“LGM 模型中缺失数据处理方法的比较及稳健性：ML 方法与 Diggle-Kenward 选择模型”。

文章没有考虑不同的非随机缺失的处理方法之间的比较，主要是由于不同的 MNAR 数据处理方法往往基于不同缺失机制的假设，如常用的 Diggle-Kenward 选择模型和模式混合模型（Pattern-Mixture Modeling）基于的缺失机制的假设就不同，基于不同假设的模型其得到的结果很难直接比较。本研究想要探讨的问题之一如果是如果缺失数据满足 D-K 模型假设的非随机缺失机制，仍采用 ML 的方法是否具有稳健性，所以本研究没有考虑更多非随机缺失机制假设下不同 MNAR 的缺失数据处理方法的比较。

不过，正如审稿专家所言，不同 MNAR 方法之间的比较这一研究问题也很有意义，这的确是本研究没有解决的问题之一。但是鉴于目前文章的重点，在修改稿中没有补充这部分的内容，但是在讨论部分进一步就这一局限性进行了说明。在讨论部分增加了“6.3 局限性”这部分的内容。详细见增加的 6.3 局限性的研究内容。

由于本部分修改内容较多，所以提交了带修订格式的文档，具体修改内容参加带修订格式的文档。

意见 2：论文的写作角度也存在较为明显的不足。一是引言部分采用了教科书式的写作，对一些概念解读过于详细，论文写作中对一些术语的解释可以更为简略。二是研究结果部分，对于图表中已经呈现的趋势，只需要做简略描述，表中已经呈现的数据，不再举例或给出重复数值。三是讨论部分应该加入与其他人研究的比较，而不是过多重复前面的研究结果。

回应：同意审稿专家的意见。在写作过程中我也感觉到有同样的问题，目前国内这一领域研究的文献不是很多，总担心一些概念解释不清楚，影响对文章的理解，这样反而使得文章很啰嗦，像写书。已经对文章整体的写法进行了比较大幅度的修改，具体修改说明如下：

(1) 引言部分：删掉了过于详细的概念解读，只列出了必要的参考文献，使得叙述尽量精炼。

(2) 结果部分：精简了结果的叙述，避免图表中的信息和文字解释部分的过多重复。详细修改见 5.1 和 5.2 两个部分，尤其是 5.2 中的叙述。

(3) 讨论部分：重新梳理了讨论部分的内容，加入了和以往研究结果的比较。

由于本部分修改内容较多，所以提交了带修订格式的文档，具体修改内容参加带修订格式的文档。

意见 3：第 6.2 节的建议部分，提出了缺失追踪数据的处理步骤，这点很实用。但与本文的内容不完全相关，可以考虑单独写一篇综述性论文，把这里所涉及的问题讲充分讲透，也很有意义。

回应：谢谢审稿专家的意见和建议。虽然说这部分的内容和本研究的研究主题不是完全一致的，但是却是实际应用者非常关心的问题，因此在修改稿中仍然保留了这部分的内容，不过将重点放在了与本研究结论相关的部分。

谢谢审稿专家的建议和鼓励，接下来将进一步阅读这一领域的文献，写一篇综述类的文章。

意见 4：向作者推荐以下三篇论文：

Zhenqiu (Laura) Lu, Zhiyong Zhang, Allan Cohen.(2013). Bayesian Methods and Model Selection for Latent Growth Curve Models with Missing Data. *New Developments in Quantitative Psychology*, pp.275-304.

Zhenqiu (Laura) Lu, Zhiyong Zhang. (2014). Robust growth mixture models with non-ignorable missingness: Models, estimation, selection, and application. *Computational Statistics & Data Analysis*.

Ke-Hai Yuan, Laura Lu. (2008). SEM with Missing Data and Unknown Population Distributions Using Two-Stage ML: Theory and Its Application. *Multivariate Behavioral Research*. 43:4, 621-652

回应：谢谢审稿专家提供的文献，已经认真阅读，并且补充了这方面的文献，在前面问题提出和讨论部分也增加了这些文献中的研究结果。同时，更新了文章参考文献。

第三轮

审稿人 1 意见：

本文经过作者的认真修改，在质量上得到了很大提升。但还可以在以下几点上进一步提高。

意见 1：关于文章的标题，作者拟用“LGM 模型中缺失数据处理方法的比较及稳健性”，而稳健性是否就是两种处理方法比较的一部分（主要比较了两种方法的估计精度和稳健性）？是否将标题定为“LGM 模型中缺失数据两种处理方法：ML 法与 D-K 选择模型的比较”更能涵盖论文的主要内容？只是建议，具体如何定标题，是作者的权利。

回应：谢谢审稿专家的建议，综合文章的内容和专家建议，现在将文章标题修改为：“LGM 模型中缺失数据处理方法的比较：ML 方法与 Diggle-Kenward 选择模型”。

意见 2：“6.1 讨论”部分新增的最后一段可以考虑放入 6.2 节中。

回应：同意专家的建议，将 6.1 中的最后一段调整到了 6.2 节中。

意见 3：除了必要的地方，建议文中可以将“Diggle-Kenward 选择模型”统一简化为“D-K 选择模型”。

回应：为了表达简练，接受专家的建议，除了文章摘要、第一次出现的地方以及图表名称，已经将文章中其他部分的“Diggle-Kenward 选择模型”简化为了“D-K 选择模型”。

意见 4：结论部分可以再斟酌，应该把最核心的结果清晰表达出来。例如，第 1 条结论实际上想比较两种方法的估计精度，却没有直接点明；第 4 条结论中“偏态程度与永久缺失比例或样本量之间通常存在一定的交互作用”中的“或”应该是“和”，其中的“通常”一词用得不好，只通过单个研究的结果不能得出通常的结论。

回应：谢谢，进一步仔细检查了这部分的内容，对部分表述不清晰的内容进行了修改。

意见 5：文字方面还存在一些问题，可以请中文好的同学或同事帮助修改一次。在摘要部分的标点使用就存在问题，如“结果表明”后面可以用冒号，同一条结论内一般不用句号。

回应：谢谢仔细审阅，已经通读全篇文字，进行了修改，并请其他同事做了审阅和修改。

意见 6：“4.3”节的公式编号有误。

回应：已经修改了公式的编号。