

严娟,郭小军,罗照盛.多维人格测验的反应与反应时联合建模分析[J].江西师范大学学报(自然科学版) 2022 46(5):453-459.
YAN Juan, GUO Xiaojun, LUO Zhaosheng. The joint modeling of response and response time in multidimensional personality tests [J].
Journal of Jiangxi Normal University(Natural Science) 2022 46(5):453-459.

文章编号:1000-5862(2022)05-0453-07

多维人格测验的反应与反应时联合建模分析

严娟¹ 郭小军¹ 罗照盛^{2*}

(1. 赣南师范大学教育科学学院,江西 赣州 341000; 2. 江西师范大学心理学院,江西 南昌 330022)

摘要:针对当前人格测验反应与反应时的数据分析模型只适用单独分析各子维度的情况,该文提出了联合单维反应时与多维反应的联合模型(JRT-M2PLM)以及联合多维反应时与多维反应的联合模型(JMRT-M2PLM),以期更准确地评估被试人格特质。在模拟研究中,发现多维联合模型JRT-M2PLM和JMRT-M2PLM的各参数均能被准确地估计,而且比单维的联合模型JRT-2PLM估计结果更优。在实际数据上,联合模型JRT-M2PLM和JMRT-M2PLM的拟合指数明显优于单维度联合模型JRT-2PLM。因此,在多维人格测验数据分析中,非常有必要考虑人格测验反应与反应时的多维性。

关键词:多维人格测验;反应时;多维模型;联合建模

中图分类号:B 841 **文献标志码:**A **DOI:**10.16357/j.cnki.issn1000-5862.2022.05.03

0 引言

随着计算机测验普遍化,测验的反应时获取也越来越便利。在能力测验中,项目反应时能反映被试在作答过程中的丰富的信息,在被试作答异常行为的识别^[1]、测验设计的改进^[2]以及潜在特质能力估计精度的提高^[3]等方面都发挥着重要的作用。在心理测验重要组成部分中,除了能力测验外,人格测验也是在心理测验中一个不可忽视的重要组成部分,因此研究者也在不断探索将人格测验的反应与反应时数据进行联合分析。

实现人格测验的反应与反应时的联合分析的一个重要的前提是理清反应时与个体特质间的关系。目前,普遍接受反应时与个体特质间的关系的理论假设是困难-距离(distance-difficulty, DD)假设。DD假设是指被试对项目的反应困难程度随着人与项目间的距离的增加而减小^[4]。P. J. Ferrando等^[5]将反应时作为反应困难指标,发现了个人与项目间的距离和项目反应时呈现负相关,进而验证了DD假设。DD假设的验证为联合分析人格测验的反应与反应时数据提供了理论基础。P. J. Ferrando等^[6]在DD假设下,将个人与项目间的距离融入对数正态反应

时模型中,实现了人格测验反应与反应时数据的联合分析。反应时在人格测验中能提供重要信息来源, J. Ranger等^[7]将反应概率模型融入对数正态反应时模型中并计算联合模型的测验信息量,结果发现考虑反应时能提高13%~17%的测验信息量。为了将DD假设从2级计分扩展到多级计分的人格测验中, Meng Xiangbin等^[8]基于反应概率角度构建了多级计分的困难-距离表达式;而J. Ranger^[9]则从项目与被试间的距离以及不同等级计分项目具有不同位置参数角度对P. J. Ferrando等^[6]和J. Ranger等^[7]的模型进行扩展,构建了二项式模型和位置模型联合分析多级人格测验的反应与反应时数据。

在多数人格测验中,常常由多个相互联系的维度构成。如中国人的大七人格测验的人格特质包括外向性、善良、情绪性、才干、人际关系、行事风格以及处事态度等7个维度,7种人格特质有机地构成了中国人的人格。这些子测验构成了一个完整的多维人格测验,此时对整个人格测验的数据分析适宜采用多维项目反应模型^[10]。但是,目前这些多维模型的应用仅仅是基于人格测验的反应数据,忽视了另一种同样能够反映被试行为的数据(即项目反应时)。现有的人格测验反应与反应时数据的联合分析方法都是单独分析被试在人格测验各维度或子测

收稿日期:2022-06-15

基金项目:江西省社会科学基金(21JY49)和江西省高校人文社会科学研究(XL20203)资助项目。

通信作者:罗照盛(1971—)男,江西吉安人,教授,博士,博士生导师,主要从事心理统计与测量研究。E-mail:luozs@126.com

验中的反应与反应时数据^[6-9],每次单独分析一个子测验或维度的方法忽略了子测验或维度之间可能存在的相关,这会增大测量误差,从而导致对被试人格特质水平做出不准确的推断^[11].此外,在能力测验中,已有研究发现联合多维反应与单维反应时的联合模型能有效提高模型参数估计准确性和模型拟合效果^[12].同时,反应时多维模型与高阶模型相应地被提出^[13],并发现多维与高阶反应与反应时的联合建模具有更强的优势^[14].

针对上述问题,在人格测验中,既有必要考虑人格测验的项目反应的多维性,又可能需要考虑反应时的多维性问题.因此,本文基于 P. J. Ferrando 等^[6]的模型,提出了分析多维人格测验的反应与反应时的联合模型,以期探究多维人格测验反应与反应时数据的合理分析方法,从而更准确地对被试人格特质进行评价.

1 反应与反应时联合模型

1.1 多维反应模型

鉴于 P. J. Ferrando 等^[6]采用 2 参数 Logistic 模型(2PLM),与之对应的多维模型为多维 2 参数 Logistic 补偿模型(M2PLM)^[10],M2PLM 为

$$\text{logit}(p(Y_{ij} = 1)) = \sum_{k=1}^K \alpha_{jk}(\theta_{ik} - b_j), \quad (1)$$

其中 $p(Y_{ij} = 1)$ 为被试 i ($i = 1, 2, \dots, N$) 在项目 j ($j = 1, \dots, m$) 上的正确作答概率, α_{jk} 为项目 j 在第 k 维度上的区分度参数, θ_{ik} 为被试 i 在第 k 维度上的潜在人格特质, b_j 为项目 j 的难度参数. 当 $K = 1$ 时,则 M2PLM 为单维 2 参数 Logistic 模型(2PLM).

1.2 反应时模型

对于人格测验的反应时数据,通常假设其服从对数正态分布^[6],因此其多维对数正态反应时模型为

$$\lg(R_{Tij}) = \beta_j - \sum_{k=1}^K \tau_{ik} + \gamma_j \delta_{ij} + \xi_{ij}, \quad \xi_{ij} \sim N(0, \sigma_j^2), \quad (2)$$

其中 $\lg(R_{Tij})$ 为被试 i ($i = 1, 2, \dots, N$) 在项目 j ($j = 1, 2, \dots, m$) 上的对数反应时, β_j 为项目 j 的时间强度参数, τ_{ik} 为被试 i 在第 k 维度上的速度潜在特质, γ_j 为个人-项目距离 δ_{ij} 的斜率参数, ξ_{ij} 为反应时残差项,其服从均值为 0、方差为 σ_j^2 正态分布,同时 $1/\sigma_j$ 为项目反应时的区分度参数.

P. J. Ferrando 等^[6]基于单维 2PLM,将个人-项目距离 δ_{ij} 定义为

$$\delta_{ij} = \sqrt{\alpha_j^2 (\theta_i - b_j)^2}, \quad (3)$$

其中 α_j 和 b_j 分别为项目 j 的区分度参数与难度参

数, θ_i 为被试 i 的潜在人格特质.

由于人格测验总是包含多个维度,同时人格测验的项目之间都是多维的,也就是一个项目只测量一种人格特质维度,不同组项目对应不同人格特质,每组项目内都是单维度的,所以,只需将不同人格特质维度的个人-项目距离 δ_{ij} 定义在同一个框架中,即可实现多维个人-项目距离 δ_{ij} 的定义.因此项目之间多维的个人-项目距离 δ_{ij} 可被定义为

$$\delta_{ij} = \sqrt{\sum_{k=1}^K \alpha_{jk}^2 (\theta_{ik} - b_j)^2}. \quad (4)$$

从式(4)可以看出:当 K 取不同数值(即特定维度)时,对应不同维度的个人-项目距离,这与式(3)单维度个人-项目距离是相同的.

对于上述模型,若个人-项目距离 δ_{ij} 为式(3)(式(4)为 $K = 1$),且反应时模型为单维对数正态模型(式(2)为 $K = 1$),则构成人格测验反应时的单维度对数正态反应时模型(ULRTM);若个人-项目距离 δ_{ij} 为式(4)且反应时模型为式(2),则构成人格测验反应时的多维对数正态反应时模型(MLRTM).

1.3 联合建模

1.3.1 单维反应时和多维反应联合模型JRT-M2PLM

基于 P. J. Ferrando 等^[6]提出的建模框架,本文首先提出基于单维反应时与多维反应模型的联合模型(JRT-M2PLM).在 JRT-M2PLM 中,反应模型 M2PLM 和反应时模型 ULRTM 构成分层模型的第 1 层,被试参数与项目参数分布构成分层模型的第 2 层.

在第 2 层被试人格特质与速度特质之间的关系上,多维人格特质与速度特质之间的关系服从多元正态分布,其可以描述为

$$(\theta_1, \theta_2, \dots, \theta_K, \pi)^T \sim \text{MVN}(\boldsymbol{\mu}_I, \boldsymbol{\Sigma}_I),$$

$$\boldsymbol{\mu}_I = \begin{pmatrix} \mu_{\theta_1} \\ \mu_{\theta_2} \\ \vdots \\ \mu_{\theta_K} \\ \mu_{\pi} \end{pmatrix}, \quad \boldsymbol{\Sigma}_I = \begin{pmatrix} \sigma_{\theta_1}^2 & \cdots & \sigma_{\theta_1\theta_K} & \sigma_{\theta_1\pi} \\ \vdots & \ddots & \vdots & \vdots \\ \sigma_{\theta_1\theta_K} & \cdots & \sigma_{\theta_K}^2 & \sigma_{\theta_K\pi} \\ \sigma_{\pi\theta_1} & \cdots & \sigma_{\pi\theta_K} & \sigma_{\pi}^2 \end{pmatrix}, \quad (5)$$

其中 $\boldsymbol{\mu}_I$ 为被试人格与速度特质的均值, $\boldsymbol{\Sigma}_I$ 为被试人格与速度特质的协方差矩阵,用于描述被试参数之间的关系.对于项目参数间关系,反应模型 M2PLM 的难度参数 b 与反应时模型 ULRTM 的时间强度参数 β 之间的关系服从二元正态分布,可以描述为

$$\begin{pmatrix} b_j \\ \beta_j \end{pmatrix} \sim \text{MVN}(\boldsymbol{\mu}_J, \boldsymbol{\Sigma}_J),$$

$$\boldsymbol{\mu}_J = \begin{pmatrix} \mu_b \\ \mu_{\beta} \end{pmatrix}, \quad \boldsymbol{\Sigma}_J = \begin{pmatrix} \sigma_b^2 & \sigma_{b\beta} \\ \sigma_{b\beta} & \sigma_{\beta}^2 \end{pmatrix}, \quad (6)$$

其中 μ_j 为项目参数均值向量; Σ_j 为项目参数的协方差矩阵,用于描述题目参数之间关系。

1.3.2 多维反应与反应时联合模型JMRT-M2PLM

在人格测验的不同维度的反应时上,可能对应不同的速度特质,因此本文提出的第2个模型为多维反应与多维反应时的多维联合模型(JMRT-M2PLM)。在JMRT-M2PLM中,第1层模型的反应模型为M2PLM(式(1)),反应时模型为MLRTM(式(2));而第2层模型也为被试参数的分布(式(5))与项目参数分布(式(6)),只是在被试参数分布中需要包含与人格特质对应的多个速度特质。

综合上述模型,人格测验往往是多维的。P. J. Ferrando等^[6]提出的单维反应与反应时联合模型JRT-2PLM只有将多维的人格测验拆分为多个单维人格才能适用。基于此,JRT-M2PLM被提出以适用于多维人格测验,并且JRT-M2PLM假设不同人格特质对应相同速度特质,即人格特质是多维的,而速度特质是单维的。然后,不同人格特质也可能对应不同的速度特质^[14],因此,一个完整的多维人格和速度特质模型被构建,即JMRT-M2PLM模型。

1.4 贝叶斯参数估计

3个联合模型JRT-2PLM、JRT-M2PLM和JMRT-M2PLM均采用贝叶斯马尔科夫链蒙特卡罗(MCMC)算法来实现参数估计,并且基于R软件(R version 4.0.2)中的R2jags包调用JAGS软件来实现。

1.4.1 模型识别 为了使模型可识别与比较,需要对一些参数进行限制,以保证参数估计的唯一性。在JRT-2PLM、JRT-M2PLM和JMRT-M2PLM等3个联合模型中,将被试人格特质参数的方差固定为1,且均值向量均为 $\mathbf{0}$ ^[12,15]。

1.4.2 先验分布 考虑到JRT-2PLM、JRT-M2PLM和JMRT-M2PLM等3个联合模型的待估参数基本一致,因此下面将统一介绍3个联合模型待估参数的先验分布。

在项目参数上,项目难度参数 b 与时间强度参数 β 服从二元正态分布(见式(6)),其中超参数 μ_b 和 μ_β 均服从正态分布($\mu_b \sim N(0,1)$, $\mu_\beta \sim N(0,1)$),协方差矩阵 Σ_j 服从二元对角矩阵 R 的逆威沙特分布($\Sigma_j \sim \text{InvWishart}(R, 2)$);模型的区分度参数 $\alpha_{kj}, 1/\sigma_j$ 均服从左截尾的正态分布 $N(0,1)I(0, \cdot)$,个人-项目距离 δ_{ij} 斜率参数 γ_j 服从右截尾正态分布($\gamma_j \sim N(0,1)I(\cdot, \rho)$)^[8,12-15]。

被试参数人格与速度特质的关系如式(5)所示,它们都服从多元正态分布。为了更好地估计协方差矩阵,对它们的协方差矩阵 Σ_i 进行Cholesky分解 $\Sigma_i = \Delta\Delta^T$,其中 Δ 为

$$\Delta = \begin{pmatrix} 1 & & & \\ \rho_{2,1} & \sqrt{1-\rho_{2,1}^2} & & \\ \vdots & \vdots & \ddots & \\ \rho_{K+k,1} & \rho_{K+k,2} & \cdots & \rho_{K+k,K+k} \end{pmatrix}, \quad (7)$$

其中 Δ^T 为 Δ 的共轭转置;其各元素的先验分布分别为: $\rho_{2,1}$ 服从均匀分布 $U(-1,1)$, $\rho_{K+k,K+k}$ 服从 γ 分布 $G(1,1)$, $\rho_{K+k,k}$ 服从正态分布 $N(0,1)$ ^[12]。

2 模拟研究

2.1 模拟设计

为了检验参数估计方法的准确性以及将多维测验拆分为单维度分析的影响,分别选择不同多维程度的JRT-M2PLM和JMRT-M2PLM为数据生成模型,然后分别基于JRT-2PLM和生成模型进行参数估计并展开模拟研究。各项目参数的生成分别为: $\alpha_{kj}, 1/\sigma_j$ 是左截尾的正态分布 $N(0,1)I(0, \cdot)$,难度参数 b 和项目时间强度参数和 β 服从均值向量分别为 $\mathbf{0}$ 和 $3E$ (其中 E 为元素全为1的向量)、协方差矩阵 $\begin{pmatrix} 1 & 0 \\ 0.25 & 0.25 \end{pmatrix}$ 的二元正态分布;在被试人格特质

维度上潜在人格特质 θ_k 和速度特质 τ_k 的协方差矩阵根据式(7)随机生成转换,均值向量固定为 $\mathbf{0}$,再通过多元正态分布式(5)生成每个被试的人格和速度特质;然后在根据JRT-M2PLM或JMRT-M2PLM生成人格测验反应与反应时数据。模拟针对测验长度($m = 20, 40$)以及被试量($N = 400, 800$)设置了 $2 \times 2 = 4$ 种条件组合,其中测验项目 m 平均分配在各个维度上,每种组合重复30次。通过计算各参数真值与估计值的 x_{MSE} (mean squared error)与平均偏差(y_{bias})来评价参数估计方法的准确性以及拆分维度分析的影响。

2.2 模拟结果

基于JRT-M2PLM和JMRT-2PLM这2个模型的模拟结果如表1和表2所示。在表1中,由于JRT-M2PLM速度特质是单维的,所以JRT-M2PLM只有1个速度特质估计结果,而JRT-2PLM将数据拆分为2个部分,所以对同一速度特质估计了2次。在项目参数上,JRT-2PLM和JRT-M2PLM各项目参数随着被试量400增加到800, x_{MSE} 值从0.02左右降低到0.01附近。在被试参数上,随着测验长度由20增加到40,JRT-2PLM的2个人格特质的 x_{MSE} 值从0.14左右降低到0.07左右,而JRT-M2PLM的 x_{MSE} 值则从0.12左右降低到0.05左右;在速度特质上,JRT-2PLM的 x_{MSE} 值则从0.10左右降低到0.07,而JRT-M2PLM

的 x_{MSE} 则从 0.06 附近降低到 0.05 左右. 此外 2 个模型的各参数的 y_{bias} 值的绝对值都小于 0.08.

表 1 基于 JRT-M2PLM 模拟的参数估计结果

| 模拟条件 | 模型 | 评价指标 | 项目参数 | | | | | 被试参数 | | | |
|-----------------------|-----------|------------|----------|--------|---------|------------|----------|------------|------------|----------|----------|
| | | | α | b | β | σ^2 | γ | θ_1 | θ_2 | τ_1 | τ_2 |
| $N = 400$ $m = 20$ | JRT-2PLM | x_{MSE} | 0.024 | 0.026 | 0.023 | 0.004 | 0.019 | 0.149 | 0.141 | 0.104 | 0.116 |
| | | y_{bias} | 0.024 | 0.005 | 0.031 | 0.002 | -0.027 | 0.005 | -0.025 | 0.018 | 0.020 |
| | JRT-M2PLM | x_{MSE} | 0.022 | 0.027 | 0.025 | 0.002 | 0.015 | 0.116 | 0.125 | 0.063 | 0.063 |
| | | y_{bias} | -0.017 | -0.010 | 0.000 | 0.003 | -0.029 | 0.003 | -0.020 | -0.024 | -0.024 |
| $N = 400$ $m = 40$ | JRT-2PLM | x_{MSE} | 0.023 | 0.028 | 0.025 | 0.002 | 0.020 | 0.063 | 0.072 | 0.055 | 0.062 |
| | | y_{bias} | 0.023 | -0.004 | -0.030 | -0.002 | -0.030 | -0.005 | 0.004 | -0.047 | -0.037 |
| | JRT-M2PLM | x_{MSE} | 0.022 | 0.026 | 0.024 | 0.002 | 0.019 | 0.051 | 0.057 | 0.036 | 0.036 |
| | | y_{bias} | -0.009 | -0.058 | -0.017 | -0.003 | -0.028 | -0.012 | -0.019 | -0.068 | -0.068 |
| $N = 800$ $m = 20$ | JRT-2PLM | x_{MSE} | 0.013 | 0.013 | 0.014 | 0.002 | 0.012 | 0.154 | 0.149 | 0.105 | 0.101 |
| | | y_{bias} | 0.013 | -0.006 | 0.020 | -0.002 | -0.009 | 0.007 | -0.004 | 0.011 | 0.010 |
| | JRT-M2PLM | x_{MSE} | 0.013 | 0.014 | 0.013 | 0.001 | 0.010 | 0.114 | 0.111 | 0.053 | 0.053 |
| | | y_{bias} | -0.015 | -0.006 | -0.012 | -0.004 | -0.015 | 0.002 | -0.011 | -0.020 | -0.020 |
| $N = 800$ $m = 40$ | JRT-2PLM | x_{MSE} | 0.012 | 0.012 | 0.032 | 0.001 | 0.009 | 0.071 | 0.075 | 0.074 | 0.073 |
| | | y_{bias} | 0.012 | 0.009 | -0.069 | -0.001 | -0.012 | 0.007 | 0.007 | -0.063 | -0.076 |
| | JRT-M2PLM | x_{MSE} | 0.013 | 0.043 | 0.012 | 0.001 | 0.009 | 0.045 | 0.050 | 0.048 | 0.048 |
| | | y_{bias} | -0.007 | -0.039 | 0.010 | -0.001 | -0.014 | -0.011 | 0.004 | -0.043 | -0.043 |

从表 2 可以看出: 在项目参数上, 随着被试量从 400 增加到 800, JRT-2PLM 和 JMRT-M2PLM 各项目参数的 x_{MSE} 值整体上从 0.02 左右降低到 0.01 左右, y_{bias} 的绝对值基本上小于 0.05, 估计较好, 但是 JMRT-M2PLM 对应项目参数估计精度要高于 JRT-2PLM. 在被试参数上, 当测验长度为 20 时,

JRT-2PLM 的 x_{MSE} 值最小为 0.110, 最大为 0.167, 而 JMRT-M2PLM 的 x_{MSE} 值最小为 0.086, 最大为 0.106; 随着测验长度增加到 40, JRT-2PLM 被试参数的 x_{MSE} 值降低到 0.06 左右, 而 JMRT-M2PLM 的 x_{MSE} 则降低到 0.05 左右, 且对应被试参数明显小于 JRT-2PLM; 在 y_{bias} 绝对值上 2 个模型普遍小于 0.10.

表 2 基于 JMRT-M2PLM 模拟的参数估计结果

| 模拟条件 | 模型 | 评价指标 | 项目参数 | | | | | 被试参数 | | | |
|-----------------------|-----------|------------|----------|--------|---------|------------|----------|------------|------------|----------|----------|
| | | | α | b | β | σ^2 | γ | θ_1 | θ_2 | τ_1 | τ_2 |
| $N = 400$ $m = 20$ | JRT-2PLM | x_{MSE} | 0.025 | 0.031 | 0.026 | 0.003 | 0.026 | 0.167 | 0.148 | 0.112 | 0.109 |
| | | y_{bias} | 0.025 | 0.016 | 0.013 | 0.001 | -0.054 | 0.009 | -0.002 | -0.001 | 0.002 |
| | JRT-M2PLM | x_{MSE} | 0.022 | 0.019 | 0.027 | 0.003 | 0.023 | 0.106 | 0.093 | 0.090 | 0.095 |
| | | y_{bias} | -0.036 | -0.004 | 0.019 | 0.000 | -0.049 | 0.013 | 0.005 | 0.000 | -0.029 |
| $N = 400$ $m = 40$ | JRT-2PLM | x_{MSE} | 0.022 | 0.026 | 0.041 | 0.002 | 0.020 | 0.080 | 0.073 | 0.084 | 0.075 |
| | | y_{bias} | 0.022 | 0.005 | -0.097 | 0.000 | -0.023 | 0.002 | 0.005 | -0.094 | -0.095 |
| | JRT-M2PLM | x_{MSE} | 0.022 | 0.050 | 0.025 | 0.002 | 0.019 | 0.059 | 0.056 | 0.088 | 0.080 |
| | | y_{bias} | -0.025 | -0.012 | 0.000 | 0.000 | -0.021 | -0.003 | 0.000 | -0.109 | -0.113 |
| $N = 800$ $m = 20$ | JRT-2PLM | x_{MSE} | 0.013 | 0.011 | 0.017 | 0.002 | 0.013 | 0.129 | 0.148 | 0.101 | 0.110 |
| | | y_{bias} | 0.013 | -0.003 | 0.029 | 0.002 | -0.022 | 0.000 | 0.007 | 0.026 | 0.013 |
| | JRT-M2PLM | x_{MSE} | 0.013 | 0.012 | 0.010 | 0.001 | 0.012 | 0.089 | 0.096 | 0.082 | 0.090 |
| | | y_{bias} | -0.006 | 0.010 | -0.002 | 0.003 | -0.020 | 0.002 | -0.001 | 0.010 | -0.007 |
| $N = 800$ $m = 40$ | JRT-2PLM | x_{MSE} | 0.012 | 0.015 | 0.014 | 0.001 | 0.010 | 0.060 | 0.059 | 0.051 | 0.062 |
| | | y_{bias} | 0.012 | 0.015 | -0.034 | 0.000 | -0.005 | 0.020 | -0.005 | 0.000 | -0.056 |
| | JRT-M2PLM | x_{MSE} | 0.012 | 0.017 | 0.014 | 0.001 | 0.010 | 0.041 | 0.039 | 0.049 | 0.058 |
| | | y_{bias} | -0.003 | -0.046 | 0.014 | 0.000 | -0.003 | 0.028 | -0.005 | -0.020 | -0.099 |

综合上述结果可以看出, 无论是基于 JRT-M2PLM, 还是基于 JMRT-M2PLM 生成的多维人格数据, JRT-M2PLM 和 JMRT-M2PLM 各参数的估计结果均普遍优于 JRT-2PLM 的估计结果, 因此多维人

格测验拆分为单维度分析会明显增加参数估计偏差. 特别是对被试参数的影响更为明显. 此外, 模拟结果也说明基于 R2jags 的 MCMC 算法对 JRT-M2PLM 和 JMRT-M2PLM 都能提供较为准确的参数估计.

3 实例数据: 艾森克人格测量问卷

3.1 数据描述与分析

这组数据来自某大学的一次人格测验. 测量工具采用的是龚耀先^[16]修订的艾森克人格测量问卷中的“神经质”(N 维度) 和“内外向”分量表(E 维度) ,每个量表选取了 11 个项目. 通过 E-prime2. 0 收集, 计时从项目出现在显示器上开始到选择下一个项目停止. 时间的记录单位为 ms ,排除无效被试, 最后有效被试 379 人.

3 个联合模型在 MCMC 估计过程中, 均使用 2 条马尔科夫链, 且包含 40 000 次迭代, 燃烧次数为 20 000, 稀疏值为 2, 最终剩余 20 000 次迭代用于参数估计. 参数拟合收敛采用潜在量尺缩减因子(potential scale reduction factor, PSRF)^[17], 各参数的 PSRF 均要求小于 1. 1, 以表示参数估计已经拟合. 2 维的区分度参数的矩阵可以表示为

$$X_{2 \times 22} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

在 JRT-2PLM、JRT-M2PLM 和 JMRT-M2PLM 等 3 模型拟合优劣比较上, 以偏差信息准则 DIC^[18] 作为评价这 2 个模型拟合优劣指标, DIC 指标是基于 MCMC 算法评价模型拟合优度最常用的指标, DIC 值越小说明该模型与实际数据越拟合.

3.2 数据分析结果

为了检验多维联合模型在多维人格测验数据分析中的优势, 通过单维度联合建模(JRT-2PLM) 和多维联合模型(JRT-M2PLM 和 JMRT-M2PLM) 在实际数据上的分析结果进行比较, 从而评价模型优劣.

在模型结构参数和模型的拟合指数上的结果如表 3 所示. 在项目结构参数上, 分别分析 E 维度和 N 维度的 JRT-2PLM 比 JRT-M2PLM 和 JMRT-M2PLM 在估计精确度上更低, 而 2 个多维联合模型估计值则非常接近, 并且 3 个联合模型的难度参数 b 与项目时间强度参数 β 的协方差的 95% 置信区间都包含了 0, 这意味着它们不存在相关. 在被试结构参数上, 3 个联合模型都发现被试人格特质与速度之间不相关, 但是 2 个多维的联合模型皆发现 E 人格特质与 N 人格特质为负相关. 在模型拟合指数上, JRT-2PLM 在 E 和 N 维度的 DIC 值之和为 $8\,572.804 + 7\,914.857 = 16\,487.661$, JRT-M2PLM 的 DIC 值为 16 089. 52, JMRT-M2PLM 的拟合指数 DIC 值为 16 018. 66, 因此从拟合指数 DIC 可以看出 JRT-M2PLM 拟合优于 JRT-2PLM, 这说明多维人格测验的不同维度的反应数据有必要联合分析; 而 JMRT-M2PLM 优于 JRT-M2PLM, 这意味着人格测验的反应时的多维性也是有必要考虑的.

表 3 不同模型的结构参数与模型拟合指数

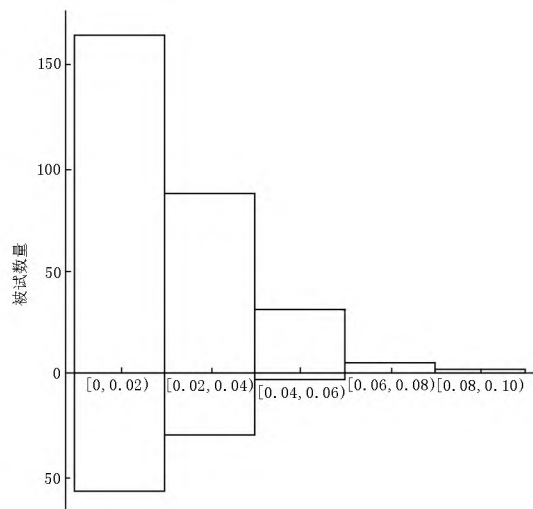
| 参数 | | JRT-2PLM | | | | JMRT-M2PLM | | JRT-M2PLM | |
|----------|-----------------------------|---------------------------------------|-------------------|--------|--------------------|------------|--------------------|------------|--------------------|
| | | M_1 | C_{I_1} | M_2 | C_{I_2} | M | C_I | M | C_I |
| 项目参数 | σ_b^2 | 0. 81 | [0. 32 , 1. 95] | 0. 70 | [0. 26 , 1. 69] | 0. 92 | [0. 48 , 1. 72] | 0. 91 | [0. 47 , 1. 72] |
| | $\sigma_{b\beta}$ | -0. 03 | [-0. 30 , 0. 22] | 0. 12 | [-0. 22 , 0. 24] | -0. 02 | [-0. 16 , 0. 11] | -0. 03 | [-0. 17 , 0. 10] |
| | σ_β^2 | 0. 16 | [0. 07 , 0. 38] | 0. 03 | [0. 03 , 0. 04] | 0. 09 | [0. 05 , 0. 16] | 0. 09 | [0. 05 , 0. 16] |
| | μ_b | 0. 14 | [-0. 42 , 0. 69] | -0. 98 | [-1. 51 , -0. 46] | -0. 39 | [-0. 80 , 0. 01] | -0. 40 | [-0. 80 , 0. 00] |
| | μ_β | 1. 38 | [1. 14 , 1. 63] | 1. 43 | [1. 20 , 1. 66] | 1. 4 | [1. 27 , 1. 52] | 1. 40 | [1. 27 , 1. 52] |
| 被试参数 | $\sigma_{\theta_1\theta_2}$ | | | | | -0. 28 | [-0. 39 , -0. 16] | -0. 29 | [-0. 41 , -0. 18] |
| | $\sigma_{\theta_1\tau_1}$ | 0. 01 | [-0. 01 , 0. 04] | | | 0. 02 | [-0. 01 , 0. 04] | 0. 00 | [-0. 02 , 0. 02] |
| | $\sigma_{\theta_1\tau_2}$ | | | | | -0. 01 | [-0. 03 , 0. 02] | 0. 01 | [-0. 01 , 0. 04] |
| | $\sigma_{\theta_2\tau_1}$ | | | | | 0. 01 | [-0. 01 , 0. 04] | | |
| | $\sigma_{\theta_2\tau_2}$ | | | 0. 01 | [-0. 02 , 0. 04] | 0. 01 | [0. 01 , -0. 02] | | |
| | $\sigma_{\tau_2\tau_1}$ | | | | | 0. 03 | [0. 03 , 0. 04] | | |
| | $\sigma_{\tau_1}^2$ | 0. 03 | [0. 02 , 0. 03] | | | 0. 03 | [0. 02 , 0. 03] | 0. 03 | [0. 03 , 0. 04] |
| | $\sigma_{\tau_1}^2$ | | | 0. 03 | [0. 03 , 0. 04] | 0. 03 | [0. 03 , 0. 04] | | |
| 拟合指数 DIC | | 8 572. 804 + 7 914. 857 = 16 487. 661 | | | | 16 018. 66 | | 16 089. 52 | |

注: M 为均值, 下标 1 和 2 分别表示维度 E 和维度 N ; C_I 为 95% 置信区间; DIC 为偏差信息准则.

模拟研究表明人格测验维度拆分分析对被试人格特质估计精度有较大影响. 考虑到 JMRT-M2PLM 模型拟合最优, 下面将通过 JRT-2PLM 在各维度上的被试的人格特质的估计标准差 S_D 与 JMRT-M2PLM 对应维度被试人格特质的估计标准差 S_D 作差, 以比较 2 个模型在人格特质上的估计精确性. 结果如图 1 和图 2 所示. 在图 1 和图 2 中, 位于 x 轴上

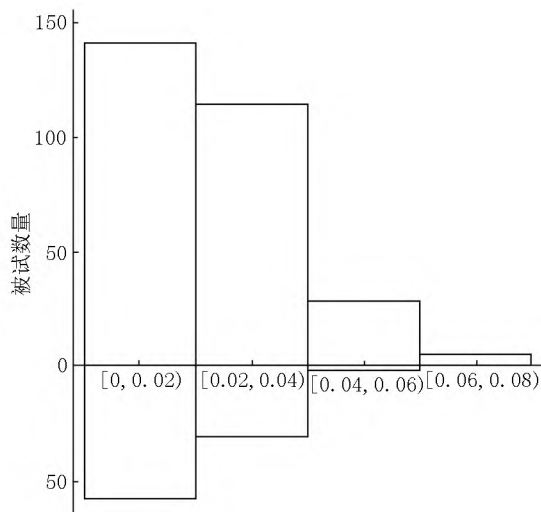
方的条形图为 JMRT-M2PLM 的被试参数的标准差 S_D 小于 JRT-2PLM 的被试参数的标准差 S_D 的被试数量(y 轴数值), x 轴的下方的条形图则为 JMRT-M2PLM 的被试参数的标准差 S_D 大于 JRT-2PLM 的被试参数的标准差 S_D 的被试数量(y 轴数值). 在 θ_1 和 θ_2 的估计标准差 S_D 上, 从图 1 和图 2 可以看出 JMRT-M2PLM 的被试标准差在绝大多数被试上要小于 JRT-2PLM 估计的结果, 特别是 2 个模型差异

较大的被试特质标准差 S_D 之差上, JMRT-M2PLM 处于绝对的优势.



JRT-M2PLM 与 M2PLM 被试参数 θ_1 估计标准差 S_D 之差

图1 JRT-M2PLM 与 M2PLM 被试参数 θ_1 比较



JRT-M2PLM 与 M2PLM 被试参数 θ_2 估计标准差 S_D 之差

图2 JRT-M2PLM 与 M2PLM 被试参数 θ_2 比较

上述结果表明基于多维人格测验反应与反应时数据的联合模型 JMRT-M2PLM 对实际数据的拟合优于单维联合模型 JRT-2PLM, 也就是说多维人格测验数据的联合分析有助于提高模型参数的估计精度与稳定性; 同时能进一步提高模型对实际数据的拟合效果.

4 讨论与总结

当前, 反应与反应时数据的联合分析模式得到了迅速的发展, 这种趋势也在人格测验研究中受到关注. 但是目前的人格测验的分析存在各种弊端, 有的只关注了人格测验的多维反应数据, 忽视了被试作答反应时^[19]; 有的则是单独分析被试在人格测验

各维度或子测验中的反应与反应时, 忽视了被试人格特质与速度特质之间的关系以及人格特质之间的关系^[5-7]. 针对这些弊端, 本文提出了联合分析人格测验的反应与反应时的多维联合模型 JRT-M2PLM 和 JMRT-M2PLM.

为了探讨单维度联合模型 JRT-2PLM 将多维人格测验的反应与反应时数据拆分为多个单维子测验或维度进行分析的弊端, 本文以 JRT-M2PLM 和 JMRT-M2PLM 为基准模型进行模拟研究, 结果发现拆分维度数据的 JRT-2PLM 在项目参数和被试参数的估计上会产生较大的偏差. 在艾森克人格测量问卷实例上, JRT-M2PLM 和 JMRT-M2PLM 相对 JRT-2PLM 在模型拟合指数和参数估计精度上仍然保持明显优势. 在项目间多维测验中, 各潜在特质之间是存在相关的, 基于单维项目反应理论模型 JRT-2PLM 的应用忽视了各人格特质之间的关系, 从而导致在参数估计和模型拟合上较差, 而多维联合模型 JRT-M2PLM 和 JMRT-M2PLM 都考虑了各人格特质间以及与速度特质间的关系, 更符合项目间多维人格测验的特征, 从而在模拟和实例上多维联合模型 JRT-M2PLM 和 JMRT-M2PLM 表现更优, 这也与已有的研究结果是相一致的^[20-22]. 因此, 在多维的人格测验反应与反应时数据分析中, 相对单维反应与反应时联合模型, 多维反应与反应时联合模型可能是一个更优的选择.

当然也有一些问题仍然有待进一步解决与完善. 如在人格与速度特质的多维性关系上, 考虑各特质之间的阶层性, 后续研究可以将反应与反应时模型扩展为高阶或双因子反应与反应时模型^[14]. 同时在反应模型上, 本文采用 2PLM 和 M2PLM 都是优势模型, 而人格测验更适用于理想点过程模型^[23], 因此后续研究可以探索在理想点过程模型下距离-困难 DD 假设的构建以及反应时与反应的联合建模.

5 参考文献

- [1] FOX J P, MARIANTI S. Person-fit statistics for joint models for accuracy and speed [J]. Journal of Educational Measurement 2017, 54(2): 243-262.
- [2] BOLSINOVA M, TIJMSTRA J. Improving precision of ability estimation: getting more from response times [J]. British Journal of Mathematical and Statistical Psychology, 2018, 71(1): 13-38.
- [3] VAN DER LINDEN W J, XIONG Xinhui. Speededness and adaptive testing [J]. Journal of Educational and Behavioral Statistics 2013, 38(4): 418-438.
- [4] KUNCEL R B, FISKE D W. Stability of response process and response [J]. Educational and Psychological Meas-

- urement, 1974, 34(4): 743-755.
- [5] FERRANDO P J. Person-item distance and response time: an empirical study in personality measurement [J]. *Psicológica* 2006, 27(1): 137-148.
- [6] FERRANDO P J, LORENZO-SEVA U. An item response theory model for incorporating response time data in binary personality items [J]. *Applied Psychological Measurement* 2007, 31(6): 525-543.
- [7] RANGER J, ORTNER T M. Assessing personality traits through response latencies using item response theory [J]. *Educational and Psychological Measurement* 2011, 71(2): 389-406.
- [8] MENG Xiangbin, TAO Jian, SHI Ningzhong. An item response model for Likert-type data that incorporates response time in personality measurements [J]. *Journal of Statistical Computation and Simulation* 2014, 84(1): 1-21.
- [9] RANGER J. Modeling responses and response times in personality tests with rating scales [J]. *Psychological Test and Assessment Modeling* 2013, 55(4): 361-382.
- [10] RECKASE M D. Multidimensional item response theory [M]. New York: Springer-Verlag 2009.
- [11] 康春花, 辛涛. 测验理论的新发展: 多维项目反应理论 [J]. *心理科学进展* 2010, 18(3): 530-536.
- [12] MAN Kaiwen, HARRING J R, JIAO Hong, et al. Joint modeling of compensatory multidimensional item responses and response times [J]. *Applied Psychological Measurement* 2019, 43(8): 639-654.
- [13] 詹沛达, JIAO Hong, MAN Kaiwen. 多维对数正态作答时间模型: 对潜在加工速度多维性的探究 [J]. *心理学报* 2020, 52(9): 1132-1142.
- [14] GUO Xiaojun, JIAO Yuyue, HUANG Zhengzheng, et al. Joint modeling of response accuracy and time in between-item multidimensional tests based on bi-factor model [J]. *Frontiers in Psychology* 2022, 13: 763959.
- [15] VAN DER LINDEN W J. A hierarchical framework for modeling speed and accuracy on test items [J]. *Psychometrika* 2007, 72(3): 287-308.
- [16] 龚耀先. 艾森克人格问卷手册 [M]. 长沙: 湖南医学院出版社, 1983.
- [17] BROOKS S P, GELMAN A. General methods for monitoring convergence of iterative simulations [J]. *Journal of Computational and Graphical Statistics* 1998, 7(4): 434-455.
- [18] SPIEGELHALTER D J, BEST N G, CARLIN B P, et al. Bayesian measures of model complexity and fit [J]. *Journal of the Royal Statistical Society* 2002, 64(4): 583-639.
- [19] 邓稳根, 戴海琦, 戴慧群. 人格测验的项目反应过程与展开模型研究述评 [J]. *心理科学进展* 2010, 18(12): 1986-1990.
- [20] YEN W M. Scaling performance assessments: strategies for managing local item dependence [J]. *Journal of Educational Measurement* 1993, 30(3): 187-213.
- [21] KAHARAMAN N. Unidimensional interpretations for multidimensional test items [J]. *Journal of Educational Measurement* 2013, 50(2): 227-246.
- [22] JEWSBURY P A, VAN RIJN P W. IRT and MIRT models for item parameter estimation with multidimensional multi-stage tests [J]. *Journal of Educational and Behavioral Statistics* 2020, 45(4): 383-402.
- [23] 邓稳根, 乐妮妮. 个人特质水平及个人-项目距离与反应时的关系 [J]. *心理学探新* 2017, 37(1): 91-96.

The Joint Modeling of Response and Response Time in Multidimensional Personality Tests

YAN Juan¹, GUO Xiaojun¹, LUO Zhaosheng^{2*}

(1. School of Education Science, Gannan Normal University, Ganzhou Jiangxi 341000, China;

2. School of Psychology, Jiangxi Normal University, Nanchang Jiangxi 330022, China)

Abstract: For the current personality test response and response time data analysis model, it is only applicable to analyze the situation of each sub-dimension separately. Therefore, a joint model of unidimensional response time and multidimensional response (JRT-M2PLM) and a joint model of multidimensional response and response time (JM-RT-M2PLM) are proposed to more accurately assess the personality traits. In the simulation study, it is found that the parameters of the multidimensional joint model JRT-M2PLM and JMRT-M2PLM can be estimated accurately and better than the unidimensional joint model JRT-2PLM. In the actual data, the fitting indices of the joint models JRT-M2PLM and JMRT-M2PLM are significantly better than the unidimensional joint model JRT-2PLM. Consequently, it is very necessary to consider the multidimensionality in the analysis of the responses and response times of the multidimensional personality test.

Key words: multidimensional personality test; response time; multidimensional model; joint modeling

(责任编辑: 冉小晓)