

认知诊断 DINA 模型研究进展

■ 张 潇 沙如雪

摘要: DINA 模型是认知诊断潜在分类模型中的一种,它是一个简单的随机连接模型,由于该模型只涉及“失误”和“猜测”两个参数,比其他模型更加简洁、灵活和易于解释,因此得到了广泛的理论和应用研究。近年来,在传统 DINA 模型的基础上,又有更加完善的 HO-DINA 模型、P-DINA 模型以及 G-DINA 模型等被提出和探讨,由此可见,DINA 模型有着良好的发展前景,同时有待相关工作者进行更深入的研究。

关键词: 认知诊断;DINA 模型;HO-DINA 模型;P-DINA 模型;G-DINA 模型

【中图分类号】 G405

【文献标识码】 A

【文章编号】 1005-8427(2013)01-0032-6

随着经典测验理论的日趋成熟和广泛应用,人们对测验的功能也提出了更高的要求。2001 年,美国正式通过法案(No Child Left Behind Act of 2001,简称为 NCLB),规定美国所有实施的测验应该给家长、老师及学生提供详细诊断信息,这个法案对认知诊断的研究起了巨大的推动作用^[1]。传统的心理和教育测验仅以测验总分或单一的能力值作为评价指标,显得过于笼统和概括,现在人们不仅要求测验能在总的能力层面进行评价,更希望能深入到内部的认知加工的层面,为此心理学家们以认知心理学和心理测量学为理论基础,开发出了具有认知诊断功能的心理计量模型(简称为认知诊断模型, Cognitive Diagnosis Model, CDM),挖掘被试作答过程和测验分数背后的知识结构、加工技能和认

知过程,做出更加准确详尽的诊断和评估,从而为采取相应的补救教学提供依据,为因材施教提供指导。

自认知诊断的概念提出以来,认知诊断模型的构建、评价和应用一直是该领域研究的热点。涂冬波和漆书青(2008)曾报告认知诊断模型已有近 60 种^[1],每种计量模型各具特点,比较成熟的有 LLTM 模型、规则空间模型、属性层次模型 DINA 模型以及融合模型(Fusion Model)等^[1,2,4]。从研究文献来看,近十年内有大量关于 DINA 模型的研究。与其他模型相比,DINA 模型采用了较简单的模型定义,仅涉及“失误”和“猜测”两参数,形式更为灵活,是一个简洁和易于解释的模型。尽管这个模型很简单,已被证明有很好的模型拟合(de la Torre and Douglas,

【作者简介】 张 潇,女,南京师范大学教育科学学院,硕士研究生(南京 210097)

沙如雪,女,南京师范大学教育科学学院,硕士研究生(南京 210097)

2004),是现在认知诊断中的一个优良模型,在实际中的应用比较广泛,体现了较好的发展前景。

1 DINA模型的原理

DINA模型(Deterministic Inputs, Noisy “and” Gate Model)适用于对二元计分项目测验进行认知诊断,该模式的创建和流行始于 Junker 和 Sijtsma (2001)的研究^[4],目前主要有 de la Torre 与 Douglas (2004)等在进行比较前沿的研究^[3,8,10,11,13,15]。

大部分CDM模型的实现都需要构建一个 Q 矩阵(K.Tatsuoka, 1985),即一个由0和1组成的 J 乘 K 的矩阵,矩阵中的 k 元素表示正确回答试题 j 所必须的属性。这里所说的属性,可以指一种技能、一个知识点或者某种加工过程。 Q 矩就是构建的一个认知矩阵,这个矩阵能够明确描述回答每个试题所需要的认知过程。建立DINA模型,要首先确定测验的认知属性,建立认知属性和项目之间的 Q 矩阵。以 de la Torre(2009)中的分数减法 Q 矩阵为例^[3],通过对小学分数减法运算的认知过程分析,得出掌握小学分数减法需要掌握以下五种认知属性:(1)基本的分数减法,(2)化简,(3)将整数部分与分数部分分开,(4)从整数部分借1,(5)将整数变为分数。例如要正确作答题目 $7\frac{3}{5} - \frac{4}{5}$,需要掌握1,3,4项属性,那么在 Q 矩阵中这个试题所对应的这一属性行向量为(1,0,1,1,0)。 $\alpha_i = \{\alpha_{ik}\}$ 表征被试属性掌握模式,也可以称为被试的知识状态, $k=1, 2, \dots, K$,当在第 k 个属性上得分为1,即 $\alpha_{ik}=1$,则说明被试掌握了认知属性 k ,得分若为0,则说明被试没有掌握属性 k 。

在DINA模型中,被试的掌握模式 α 和 Q 矩阵产生了一个潜在的作答向量(a latent response vector) $\eta_{ij} = \{\eta_{ij}\}$, η_{ij} 是 $\{\alpha_{ik}\}$ 和 $\{q_{ik}\}$ 的函数, $\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{ik}}$,在这个公式中,若被试 i 掌握了项目 j 考核的所有属性,则 $\eta_{ij}=1$;若被试 i 至少有一个项目 j

考核的属性未掌握,则 $\eta_{ij}=0$ 。如果排除失误和随机因素,对某个试题的正确作答概率只有0或1两种可能,被试的反应也仅取决于 α 和这个试题 Q 矩阵的交互作用。然而,在潜在的作答过程中肯定有猜测等成分存在,所以这个潜在作答向量仅代表了一种理想的反应模式。混入这个过程“noise”就是指“失误”和“猜测”这两个参数,也就是说,掌握了某个试题所要求的全部属性的被试可能因为失误将试题答错,而那些缺少至少一项某个试题所要求的属性的被试却有可能通过猜测正确回答了这个问题。

在DINA模型中,试题 j 的失误和猜测参数用分别 s_j 和 g_j 来定义。下面公式(1)中的 Y_{ij} 指被试在试题 i 上的反应,回答正确,则其值为1,反之值为0。

$$\begin{aligned} s_j &= P(Y_{ij}=0 | \eta_{ij}=1), \\ g_j &= P(Y_{ij}=1 | \eta_{ij}=0) \end{aligned} \quad (1)$$

因此掌握模式为 α_i 的被试 i 正确回答试题 j 的概率可用下面的公式计算:

$$P(Y_{ij}=1 | \alpha_i) = (1-s_j) \eta_{ij} g_j^{1-\eta_{ij}} \quad (2)$$

在项目反应理论局部独立的假设下,DINA模型的似然函数为:

$$L(s, g, \alpha) = \prod_{i=1}^N \prod_{j=1}^m \left\{ p_j(\alpha_i)^{Y_{ij}} [1-p_j(\alpha_i)]^{1-Y_{ij}} \right\} \quad (3)$$

de la Torre & Douglas (2004)还指出^[3],这里所说的猜测并不是指完全随机作答,而是包含了没有反映在 Q 矩阵中的其他解题策略。例如,如果一个试题可以用不同的认知属性来解决,这些属性未被包含在 Q 矩阵中,那么具备这些属性的被试就能利用不同的策略系统地去解决问题,而并非是通过猜测答对。

2 DINA模型的改进

2.1 HO-DINA模型

在认知诊断中, α 作为一种知识状态,它的每一个属性元素是对某种特定的规则或信息的掌握指

标。De la Torre & Douglas(2004)^[5]认为一个能评估这些属性掌握情况的模型应该被假定这些属性与一种或几种更高层的一般智力或一般能力相联系,那些具备了高层能力(θ)的被试更有可能获得了测试项目要求的认知属性。de la Torre 与 Douglas(2004)把包含了 θ 的 DINA 模式就称为 higher-order DINA 模式(HO-DINA)。在传统 DINA 模型基础上,假设在给定 θ 的前提下认知属性 α_i 条件独立,则关系式可用下列式子(4)(5)表示

$$P(\alpha|\theta) = \prod_{k=1}^K P(\alpha_k|\theta), \quad (4)$$

$$p(\alpha_k|\theta) = \frac{\exp[1.7\lambda_{1k}(\theta - \lambda_{0k})]}{1 + \exp[1.7\lambda_{1k}(\theta - \lambda_{0k})]} \quad (5)$$

由公式可以看出,这个模型是一个包含潜在协变量 θ 的逻辑回归模型。它与试题反应理论(IRT)中的单参数对数模型相似,类似于一个以认知属性为项目的更高层次上的项目反应模型, λ_1 是指属性 k 的定位参数,它类似于属性掌握的难度参数, λ_{0k} 指斜率。高阶模型建立在传统模型基础上,增加了较“属性”更高阶的“能力”参数,降低了技能组合的数量,不仅能描述被试的总体水平(θ),还能描述被试对属性的掌握情况以及被试掌握属性与能力的关系,从而能够提供更丰富的信息^[5]。

不管是以上提到的 DINA 模型,还是改进的 HO-DINA 模型,都是一种单策略模型,每一个项目都对应唯一的认知属性向量。而实际上,可以用多种策略对一个项目进行解答,因此,对于一个项目可以构建多个认知属性向量。多策略 DINA 模型^[6]在单策略的 DINA 模型上直接进行了扩展,假设每个项目包含了能够充分解决问题的 M 种不同的策略,每种策略是 K 属性向量集合中的一个子集,可以用 M 个不同的矩阵表示, Q_1, Q_2, \dots, Q_M 。对于被试 i 和项目 j ,重新定义, $\eta_{ijm} = \prod_{k=1}^K \alpha_{ik}^{q_{jkm}}$, $m=1, 2, 3, \dots, M$, Q_{jkm} 表示 Q_M 的第 j 行第 k 列, n_{ijm} 表示被试 i 能否

用第 m 种策略解答项目 j 。检测是否至少有一种策略满足条件,一旦 η_{ij} 确定之后,就可以用 DINA 模型中的公式进行计算了。这个模型的定义很简单,它假设 s 和 g 的值对于任何策略都是一样的,因此应用此模型的前提条件必须要求不同策略的难度要相等,这显然是有局限性的,因此如何能够允许 s 和 g 参数在不同的策略条件下变化对于多策略模型是否具有适用性至关重要,对于这个问题还需要更进一步研究。

2.2 P-DINA模型

现在已开发的认知诊断的 60 多种模型基本上都是适用于 0-1 评分数据,目前应用广泛的 DINA 模型模型也仅适用于 0-1 评分数据,而实际情境中,通过一般教育与心理测验所得的数据基本都是多级评分数据。涂冬波,蔡艳等(2010)对 DINA 模型进行拓广,开发出同时适合 0-1 评分与多级评分数据的 DINA 模型(简记为 P-DINA 模型)^[7],并采用 MCMC 算法实现其参数估计,同时对其性能进行研究,为认知诊断在实际中的应用提供一种新模型。

P-DINA 模型 DINA 模型基于 Samejima (1997)的等级反应模型中的累积类别反应函数思想,它的概率反应函数为:

$$P(Y_{ij} = t | \alpha_i) = P^*(Y_{ij} = t | \alpha_i) - P^*(Y_{ij} = t + 1 | \alpha_i)$$

$$P^*(Y_{ij} = t | \alpha_i) = (1 - s_{ij})^{\eta_{ij}} g_{ji}^{1 - \eta_{ij}} \quad (6)$$

根据项目反应理论局部独立性假设,可得出 P-DINA 模型的似然函数如下:

$$L(s, g, \alpha) = \prod_{i=1}^N \prod_{j=1}^m \prod_{t=1}^{mf_j} p_{ijt}^{u_{ijt}} \quad (7)$$

P-DINA 模型下,估计的参数更多,除了 0-1 评分模型下要估计的 α 参数外,还要估计 mf 个失误参数 s 和 mf 个猜测参数 g ,涂冬波等人采用当前国际比较先进的 MCMC 算法来实现了此模型参数估计。经蒙特卡洛模拟实验研究表明,P-DINA 模型具有良好的性能,参数估计的精度较高,若想保证属性

模式判断率在80%以上,建议诊断的属性个数不宜超过7个。

2.3 G-DINA 模型

DINA 模式的一般化模型 (Generalized DINA Model, G-DINA) 由 de la Torre (2011) 提出^[8], 比起 DINA 模型, G-DINA 放松了对条件的限制, 它的饱和模型与基于其他连接函数的认知诊断的一般模型是等值的。在合适的限制条件下, 其他一些常用的认知诊断模型都可以看成一般模型的特例。可以说, G-DINA 模型为 CDM 提供了一个一般性的框架模型, 用方程式表示:

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{jk} + \sum_{k=k+1}^{K_j^*} \sum_{k'=1}^{K_j^*-1} \delta_{jkk'} \alpha_{kk'} \cdots + \delta_{j12 \cdots k_j^*} \sum_{k=1}^{K_j^*} \alpha_{kk} \quad (8)$$

δ_{j0} 为试题 j 的截距, 表示没有掌握属性而正确回答的基线概率。 δ_{jk} 为对 a_k 的主要影响, 表示仅掌握了一种属性引起正确回答概率的变化。 $\delta_{jkk'}$ 表示 α_k 和 $\alpha_{k'}$ 交互影响。 $\delta_{j12 \cdots k_j^*}$ 表示 $\alpha_1, \alpha_2, \dots, \alpha_{k_j^*}$ 的交互影响。当 δ_{j0} 和 $\delta_{j12 \cdots k_j^*}$ 不为 0, 而 $\delta_{jk}, \delta_{jkk'}$ 均为 0 时, 就是 DINA 模型了, 所以 DINA 模型是 G-DINA 模型的特例。de la Torre 对 G-DINA 的估计是采用 EM 算法, 程序码是使用高级编程语言 Ox (Doornik, 2003) 编写的, de la Torre 指出, 如果参数估计程序编码能够由更专业的程序人员用低级的语言编写出来, 这将使得 G-DINA 模型在大规模数据处理上更加适用。

3 DINA 模型的特点

DINA 模型是一个随机连接模型 (stochastic conjunctive model), 之所以“随机”, 因为被试答对项目并非由被试掌握的属性唯一决定, 还加入了 s_j 和 g_j , 也就是说掌握了所有属性不能保证一定答对, 而未掌握所考属性也不一定答错。这里的“连接”意

味着所有的属性具非补偿性, 被试某一属性没有掌握是不能通过其它属性来补偿, 只有一个属性没有掌握和所有属性都没有掌握是一样的。DINA 模型与研究得较广泛的空间规则模型一样, 都要建立 Q 矩阵, 不同的是, DINA 模型不需要考虑层次间的属性关系, 不是通过分类或判别方法求得被试的掌握模式, 而是通过 MCMC 方法把它作为待估参数进行估计, 这些参数包括被试的掌握模式以及 s_j, g_j 。关于参数 s_j 和 g_j 有两个值得注意的地方: 首先模型应满足 $1-s_j > g_j$; 另外在 DINA 模型参数估计中, 要求 s_j 和 g_j 值不能过大, 一般是不大于 0.4 为界限, 如果超过 0.4, 可能说明 Q 矩阵界定的不好, 或者是被试使用的其他解题策略, 模型不拟合资料等^[9]。

4 DINA 模型的相关研究

4.1 模型的性能及改进研究

DINA 模型在国外已经进行了相对广泛的理论研究, 国内对这个模型的研究虽然不多, 但是也有学者专门对此进行过探讨。一方面, 很多研究着重于对该模型的改进方面, de la Torre 与 Douglas (2004) 认为受试者的能力是多元的, 应与试题难度、区分度相对应, 并以此为基础提出高阶的 DINA 模型 (Higher-order DINA 模型), 采用了 MCMC 方法对其模型进行了参数估计^[5]; de la Torre (2009) 还针对选择题型, 提出 multiple-choice DINA 的模式, 试图从选项中获得更多的诊断讯息, 达到更精准的估计^[10]。江西师范大学的涂冬波 (2010) 对简单的二级评分模型进行改进, 提出了 DINA 的多级评分模型^[7]。de la Torre (2011) 又在传统的 DINA 模型基础上放松对条件的限制, 提出 DINA 的一般化模型 (G-DINA 模型)^[8]; 另一方面, 很多学者对于模型的性能也进行了大量研究, de la Torre 与 Douglas (2004) 探讨了 DINA 与 LLM 模型的比较, 得出用 MCMC 法对 DINA 的参数估计精准度较高的结论; Zhang W.M. (2006) 研究了 DINA 模型下的 DIF 检测

方法^[9];Rupp & Templin (2008)研究了 Q 矩阵的不完整性对 DINA 模型诊断结果的影响^[11];Rupp & Templin (2007)和 Cheng (2008)的研究表明该模型具有较高的判准率;另外,在模型的算法方面,de la Torre (2009)详述了 DINA 参数估计的方法,如 joint maximum likelihood estimation 及 marginalized maximum likelihood estimation 等,降低 MCMC 参数估计的时间^[3]。由此可见,有关 DINA 模型的理论研究已经比较成熟,针对其缺陷改进的模型也均通过模拟和实证数据资料等证明了其诊断的可靠性以及参数估计方法的精确度,这都为 DINA 模型的进一步实际应用打下了基础。

4.2 模型的应用研究

近年来,在 DINA 模型的应用研究方面也取得了一定的进展。在测验编制和计算机自适应测验方面,Henson & Douglas (2005)提出了基于 K-L 信息量(Kullback-Leibler Information)在 DINA 模型下挑选认知诊断测验的项目或组卷研究;Finkelman & Roussos (2009)提出利用遗传算法进行自动编制认知诊断模型测验;Cheng Y. & Chang H. (2009)在 DINA 模型下进行了认知诊断的计算机自适应测验(CD-CAT)研究^[13]。其他方面,Junker & Sijtsma (2001)用该模型来研究传递推理,Templin & Henson (2006)将这个模型用于病理性赌博的研究^[2]。此外,此模型在研究儿童心理发展和学业能力测验诊断中应用较多。国内涂冬波(2009)用这个模型对项目自动生成的小学儿童数学问题解决认知诊断 CAT 编制进行了研究^[9]。吴芳菲(2009)用该模型对六年级学生数学评价能力评价的研究,陈艳梅(2009)用该模型对初中三年级学生阅读能力进行评价研究等。在教育测验中,应用 DINA 模型进行认知诊断,不但可以向家长、老师和学生等各方面提供学生的掌握模式,还能根据对试题猜对或失误的概率指导测验的编制,提高试卷质量,从而更科学准确地对学生的学业能力进行评估。

5 认知诊断 DINA 模型的应用前景

认知诊断作为新一代测验理论,现在已被认为是心理测量理论的核心,在过去的几年中,吸引了大批认知心理学家、统计测量学家及各专门学科专家的关注,已经成为测量研究中的热门领域,包括不断提出不同的模型以及用 EM 和 MCMC 方法对模型参数进行估计的大量研究,并在教育实践中进行了验证和初步应用。在当今教育改革的形势下,认知诊断能够为教学评价、学习障碍诊断等提供更加具体更加准确的信息,使教学目标和相应补救措施的采取更具针对性,因而已经被成功应用到了大规模的教育测验中。比起其他 CMD 模型,DINA 模型不需要考虑层次间的属性关系,参数易于识别,也易于解释,是一种优良的认知诊断模型,逐渐得到了越来越多的关注和研究。但是该模型定义过于简单,对心理过程的描述不够充分,这在一定程度上限制了它在实际测验中的应用,近年来研究者在传统 DINA 基础上模型上又进行了改进和扩展,出现了它的高阶模型,多级评分模型和一般模型等,尤其在计算机自适应测验相结合方面取得了进步,相信随着理论探索的不断加深,其应用前景也将更加广阔,也必将在当前的教育和心理测量研究中发挥越来越重要的作用。

参考文献

- [1] 涂冬波,蔡艳,戴海崎,漆书青. 现代测量理论下四大认知诊断模型述评[J]. 心理学探新,2008,28(2):64-68.
- [2] 甘媛源,余嘉元. 心理测量理论的新进展:潜在分类模型[J]. 宁波大学学报,2008,30(6):61-64.
- [3] de la Torre, J. DINA model and parameter estimation: A didactic [J]. Journal of Educational and Behavioral Statistics, 2009,34: 115-130.
- [4] Junker B W, & Sijtsma K. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory [J]. Applied Psychological Measurement, 2001,25(3): 258-272.
- [5] de la Torre, J., & Douglas, J. Higher-order latent trait models for cognitive diagnosis[J]. Psychometrika. 2004,69(3):333-339.

- [6] de la Torre, J., & Douglas, J. Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data [J]. *Psychometrika*, 2008, 73(4): 595–601.
- [7] 涂冬波, 蔡艳, 戴海琦, 丁树良. 一种多级评分的认知诊断模型: P-DINA 模型的开发[J]. *心理学报*, 2010, 42(10): 1011–1013.
- [8] de la Torre, J. The generalized DINA model framework[J]. *Psychometrika*, 2011, 76(2): 179–183.
- [9] 涂冬波. 项目自动生成的小学儿童数学问题解决认知诊断CAT 编制[D]. 江西师范大学博士学位论文, 2009: 37–125.
- [10] de la Torre, J. A cognitive diagnosis model for cognitively-based multiple-choice options[J]. *Applied Psychological Measurement*, 2009, 33: 163–183.
- [11] Rupp, A., & Templin, J. The effects of q-matrix misspecification on parameter Estimates and classification accuracy in the DINA model[J]. *Educational and Psychological Measurement*, 2008, 68(1): 78–96.
- [12] de la Torre, J., & Lee, A. A note on the invariance of DINA model parameters. *Journal of Measurements*, 2010, 47: 115–127.
- [13] Cheng, Y., & Chang, H. When cognitive diagnosis meets computerized adaptive testing: CD-CAT[J]. *Psychometrika*, 2009, 74(4): 619–632.

Research Advance in DINA Model of Cognitive Diagnosis

ZHANG Xiao and SHA Ruxue

Abstract: The DINA model is one of the cognitive diagnosis model belonging to the latent class ones, and it is a simple stochastic conjunctive model. As only two parameters, “slip” and “guess”, be involved, this model is more concise, flexible, and easy to explain, leading it to a widely theoretical and applied research. In recent years, based on the traditional DINA model, there are more comprehensive HO-DINA model, P-DINA model and G-DINA model being proposed and discussed. In conclusion, DINA model has good development prospects, and need more in-depth explore by relevant experts.

Keywords: DINA Model; HO-DINA Model; P-DINA Model; G-DINA Model

(责任编辑 吴四伍)