

文章编号: 1000-5862(2018)01-0067-07

一种非参数的多策略方法: 多策略的海明距离判别法

李元白, 曾平飞, 杨亚坤, 康春花*

(浙江师范大学教师教育学院, 浙江 金华 321004)

摘要: 将单策略汉明距离判别法(SS-HDD)拓展至多策略情境(MS-HDD), 并通过多因素的模拟设计及与已有研究的比较, 考察其精确性与适宜性. 研究结果表明: 1) MS-HDD 构建思路简洁, 无需前提假设和参数估计, 计算简便; 2) 单策略情境: MS-HDD 与 SS-HDD 具有等同的判别效果, 且判准率均高于相同条件下的广义距离模型(GDD)和多策略广义距离模型(MSCD); 3) 多策略情境: MS-HDD 具有与 MSCD 相当的 PMR 和 SMR, MS-HDD 可适于多策略测验情境.

关键词: 认知诊断; 多策略; 海明距离判别法; 判准率

中图分类号: B 841.7 文献标志码: A DOI: 10.16357/j.cnki.issn1000-5862.2018.01.12

0 引言

在心理与教育测量学中, 认知诊断评估(Cognitive Diagnosis Assessment, CDA) 是对被试问题解决时认知结构或技能的优势与不足所进行的诊断与评估. 目前测量学者已经开发出了 100 多种认知诊断计量模型^[1]. 然而, 已有的这些模型都是针对单策略测验情景提出的, 即模型均假设所有被试解答所有问题时, 均采用了同一种加工或解题策略. 但在实践中, 由于被试知识结构或认知特质的差异性, 以及测试形式和题型的多样化, 面对不同问题甚至同一问题时, 他们有可能采用不同的加工方式或解题策略. 因此, 由于忽略了被试解题策略的多样性以及差异性^[2], 单一策略认知诊断模型在实践中的应用受限.

C. Y. Chiu 等^[3] 通过采用多个 Q 矩阵来表征多种解题策略的思想, 构建了多策略 DINA 模型(Multiple Strategies-DINA, MS-DINA). 涂冬波等^[2] 将广义距离判别法(GDD)拓展至多策略情境, 祝玉芳等^[4] 进一步将其拓展至多级评分多策略情境, 以吻合评分多样化的实践需求. 然而, 这些多策略方法都是对参数模型的拓展. 参数模型能实现对模型中各参数的精细估计, 但同时也存在一些不足, 如需要编制复杂的参数估计程序, 属性数目太多或样本容量较小时参数估计会不准确, 从而导致判准率降低. 为此, 非参数方法的探索逐步崭露头角, 如 C. Y. Chiu 等^[3]

和康春花等^[5] 提出的聚类分析方法, C. Y. Chiu 等^[6] 和罗照盛等^[8] 提出的基于理想反应模式(Ideal Response Pattern, IRP) 与观察反应模式(Observed Response Pattern, ORP) 之间的海明距离判别法(Hamming Distance Discrimination, HDD) 等. 非参数方法具有无需前提假设、容易理解、操作简单、耗时少、可适于小样本课堂评估等特点^[5, 8], 使其成为近几年研究的关注点.

在诸多非参数方法中, 又以 HDD 最为简洁, 因其只要获取学生的作答数据, 进行简单的加减运算即可实现对知识状态的诊断分类. 从思想和方法上来说, 易于非专业人士和一线教师理解, 为 CDA 真正走向实践提供更多的可能. 为此, 本研究拟基于前人多策略扩展的思想和方法, 将 HDD 这种简便易行的单策略海明距离判别法(Single-Strategies Hamming Distance Discrimination, SS-HDD) 拓展为多策略的海明距离判别法(Multiple-Strategies Hamming Distance Discrimination, MS-HDD), 并通过多因素条件下的模拟研究考察该方法的稳定性与准确性.

1 多策略 DINA 模型和多策略广义距离判别法

1.1 多策略 DINA 模型

de la Torre 等^[9] 提出的多策略 DINA 模型假定

收稿日期: 2017-05-17

基金项目: 教育部人文社会科学研究规划基金(16YJA190002) 资助项目.

通信作者: 康春花(1974-), 女, 江西弋阳人, 副教授, 博士, 主要从事心理测量与评价方面的研究. E-mail: akang@zjnu.cn

测验中的每道题目均可采用 M 种不同的策略进行解决,每种策略所需的认知属性可各不相同,则 M 种策略对应 M 个 Q 矩阵.并且,假设每种策略的难度相同,则在不同策略下,项目的猜测参数(g)和失误参数(s)均保持不变.由此,第 i 个被试使用第 m 种策略作答第 j 题的潜在反应变量可以定义为 $\eta_{ijm} = \prod_{k=1}^K \alpha_{ik}^{q_{jkm}}$,其中 q_{jkm} 表示应用第 m 种策略解决第 j 题时,是否考查了第 k 个属性, α_{ik} 表示被试 i 是否掌握了属性 k .取 M 个潜在反应变量的最大值作为被试 i 在项目 j 上的潜在反应,其判别方式为 $\eta_{ij} = \max(\eta_{ij1}, \eta_{ij2}, \dots, \eta_{ijM})$, $P(Y_{ij} = 1 | \alpha_i) = P(Y_{ij} = 1 | \eta_{ij}) = (1 - s_j)^{\eta_{ij} g_j^{1-\eta_{ij}}}$.

图1清晰地给出了MS-DINA的思维导图^[9].在给定的被试知识状态 α_i 的情况下,由于不同策略下 Q 矩阵的不同,所以不同策略的 Q 矩阵对应第 j 题则会有 M 种 q_{jm} ,分别对应了 M 种潜在反应变量,从其中选择最大的作为该被试最终的潜在反应,然后依据猜测参数和失误参数结合其观测分数便可得该名被试的属性掌握模式.

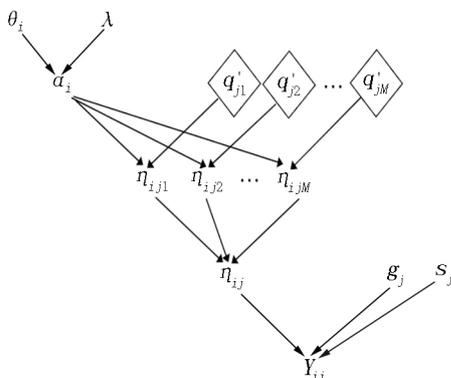


图1 MS-DINA 思维导图

1.2 多策略广义距离判别法

涂冬波等^[2]基于MS-DINA的思想和GDD提出了多策略广义距离判别法,并命名为MSCD.该方法的诊断过程是:(i)计算每种策略下被试的IRP及其对应的属性掌握模式;(ii)依据被试不同的ORP,采用IRT中2PLM估计被试的能力参数;(iii)采用孙佳楠等^[10]的GDD方法计算被试ORP到每种策略下的各个IRP的广义距离.公式定义为

$$d(X_i, I_i^m) = \sum_j^J d(X_{ij}, I_{j(i)}^m), \quad (1)$$

$d(X_{ij}, I_{j(i)}^m) = |X_{ij} - I_{j(i)}^m| P_j(\theta_i)^{X_{ij}} (1 - P_j(\theta_i))^{1-X_{ij}}$, (2) 其中 J 为项目个数, X_i 为被试 i 的 ORP, I_i^m 表示第 m 种策略下第 t 种 IRP, $d(X_{ij}, I_{j(i)}^m)$ 表示被试 i 在项目 j 上的观察反应与第 m 种策略下理想反应间的广义距

离. $d(X_i, I_i^m)$ 指被试 i 的 ORP 与第 m 种策略下第 t 种 IRP 之间的广义距离, θ_i 是由被试 ORP 估计的能力值^[11].涂冬波等^[2]研究发现, MSCD 的判准率与策略数目没有关系,是一种被试间多策略方法,可以诊断出被试在一份测验上倾向于适于何种策略^[12].

2 海明距离判别法的多策略拓展

2.1 海明距离判别法

C. Y. Chiu 等^[6]提出了一种基于距离判别的非参数认知诊断方法——海明距离判别法(HDD).而后,罗照盛等^[7]详细地阐述了HDD的判别功效,并将其中的R方法扩展至B方法,即当某个ORP与多个IRP具有同样最小值的情况下,R方法将随机选择某一KS判别给被试,B方法则扩展至采用Bayes判别法对于多个最小值进行筛选.该距离判

别法的公式为 $d_h(y_i, \eta_i) = \sum_{j=1}^J |y_{ij} - \eta_{ij}| \eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}^*}$,其中 η_{ij} 为第 t 种属性掌握模式在第 j 题上作答得出的理想反应模式, $\eta_{ij} = 1$ 说明被试 i 具有第 j 题所考查的所有属性,否则代表至少有一项属性不被掌握; y_i 为被试的实际作答反应.选取与实际作答反应距离最小的理想反应向量作为其被试的IRP,并依据在一定条件下可达阵可以使得知识状态和理想反应模式——对应的结论^[12-13],由该IRP推测该名被试的理想掌握模式即被试的知识状态.可见,HDD的思想简单易懂、无需参数估计,且具有较高的判准率^[7].因此,能否将HDD拓展至多策略情境,并考察其稳定性和准确性是本研究拟关注的问题.

2.2 多策略海明距离判别法

2.2.1 被试的策略转换 被试的策略转换可以分为被试间策略转换和混合策略转换.被试间策略转换即不同被试可能具有不同的解题策略,但同一名被试在一套题目上的策略相同.混合策略则指不同被试可能采用不同的解题策略,甚至同一被试在不同项目也有可能采用不同策略.研究表明,在策略的使用方面,较常见的是被试间策略转换^[4,14],而关于混合策略使用的研究结果则较不理想^[11].

2.2.2 多策略海明距离判别法的构建 基于已有关于被试策略转换及多策略拓展的思路,本研究将单策略海明距离判别法(SS-HDD)拓展至多策略海明距离判别法(MS-HDD),其表达式为 $d_h(y_i, \eta_{im}) =$

$$\sum_{j=1}^J |y_{ij} - \eta_{ijm}| \eta_{ijm} = \prod_{k=1}^K \alpha_{ik}^{q_{jkm}} \mathcal{C} = \arg \min\{d_h(y_i, \eta_{im}) \mid m = 1, 2, 3, \dots, M\}$$
 其中 $m = 1, 2, 3, \dots, M$ 表示在所有的 M 种策略中,使用第 m 种策略作答的情况。 y_i 表示第 i 名被试在所有题目上的观察反应模式。 η_{ijm} 表示在某一特定的第 m 种策略下,对于第 t 种属性掌握模式在项目 j 上的理想反应模式。 y_{ij} 表示被试 i 作答 j 题所得到的观察反应,计算 y_i 与所有策略可能的理想反应向量之间的绝对值的总和,将这种方法称为多策略的海明距离。 \mathcal{C} 表示分类的决策,即使得 $d_h(y_{ij}, \eta_{ijm})$ 最小的海明距离所对应的属性掌握模式作为该名被试 i 的属性掌握模式,而使得该距离值最小的 m 为该名被试的解题策略。当 $m = 1$,即测验情境为单策略时,MS-HDD 就等同于 SS-HDD。

2.2.3 MS-HDD 的判别思路 由于在确保测验 Q 矩阵中包含 R 矩阵,则可以实现属性掌握模式也即知识状态与 IRP 一一对应^[12-13]。所以,本文在保证测验 Q 矩阵中包含 R 矩阵的条件下,依据 HDD 的思想,寻找与某名被试的 ORP 距离最小的 IRP,将该 IRP 所对应的 KS 判别给该名被试^[7]。据此,MS-HDD 的判别思路为:(i) 先构造出不同策略下的 Q 矩阵,并依据该 Q 矩阵计算不同策略下的 IRP 及属性掌握模式;(ii) 根据涂冬波等^[2]的思路,即选取使与待测被试 ORP 距离最小的策略下的 IRP 对应的 KS 判别给该被试,从而完成多策略下的判别。因此该步骤根据被试的 ORP,将不同策略下所有的 IRP 与该名被试的 ORP 进行距离比较,获得与该名被试 ORP 距离最小的 IRP,并将该 IRP 对应的属性掌握模式判别给该被试。当存在一个 ORP 对应多个 IRP 的情况时,本研究采用 R 方法对被试进行归类判别;(iii) 以此类推,直至将所有被试都判别到其所属的属性掌握模式。

3 MS-HDD 在多策略情境中的适用性

3.1 研究目的

通过与 SS-HDD 以及已有多策略方法的比较,考察 MS-HDD 在多策略情境下的适用性。

3.2 研究假设

MS-HDD 具有较高的判准率。在单策略情境下,MS-HDD 的判准率应与 SS-HDD 一致或差异充分小;而在多策略情境下,MS-HDD 的判准率应高于 SS-HDD。

3.3 研究设计

3.3.1 多策略 Q 矩阵的设定 为便于与已有多策略方法结果的比较,采用涂冬波等^[2]的多策略 Q 矩阵,见表 1。涂冬波等^[2]基于 de la Torre 等^[9]的多策略 Q 矩阵,依据 Q 矩阵包含 R 矩阵,可实现理想掌握模式和 IRP 一一对应的原则^[12-13],在原有多策略 Q 矩阵基础上加入了 R 矩阵,构造出如表 1 所示的多策略 Q 矩阵。

表 1 多策略 Q 矩阵

项目	属性									
	策略 A					策略 B				
	A1	A2	A3	A4	A5	A3	A4	A5	A6	A7
1	1	1	0	0	0	0	1	0	1	1
2	1	0	1	0	0	0	0	1	1	1
3	1	0	0	1	0	0	1	1	0	1
4	1	0	0	1	0	0	1	1	0	1
5	0	1	1	0	0	1	0	0	1	1
6	0	1	0	1	0	1	0	0	1	1
7	0	1	0	0	1	1	1	0	1	0
8	0	0	1	1	0	1	0	1	0	1
9	0	0	1	0	1	1	0	1	1	0
10	0	0	0	1	1	1	1	1	0	0
11	1	1	1	0	0	0	0	0	1	1
12	1	1	0	1	0	0	1	0	0	1
13	1	1	0	0	1	0	1	0	1	0
14	1	0	1	1	0	0	0	1	0	1
15	1	0	1	0	1	0	0	1	1	0
16	1	0	0	1	1	0	1	1	0	0
17	0	1	1	1	0	1	0	0	0	1
18	0	1	1	0	1	1	0	0	1	0
19	0	1	0	1	1	1	1	0	0	0
20	0	0	1	1	1	1	0	1	0	0
21	1	0	0	0	0	0	1	1	0	0
22	0	1	0	0	0	1	0	0	1	0
23	0	0	1	0	0	1	0	0	1	0
24	0	0	0	1	0	1	1	0	0	0
25	0	0	0	0	1	1	0	1	0	0
26	1	1	0	0	0	1	0	0	0	0
27	1	0	1	0	0	0	1	0	0	0
28	1	0	0	1	0	0	0	1	0	0
29	1	0	0	0	1	0	0	0	1	0
30	0	1	1	0	0	0	0	0	0	1

3.3.2 实验设计 研究假定属性之间无补偿,且各属性之间独立。本研究为 $2 \times 3 \times 2$ 的 3 因素混合设计,测验情境(单策略情境、多策略情境)和随机失误率(2%、5%、10%)为被试间变量,判别方法(SS-HDD、MS-HDD)为被试内变量。每种条件模拟 30 次以减少误差,因变量为 30 次的判准率均值。

3.3.3 测验情境的设计 如表1所示.策略A与策略B中A3~A5的3个属性是共同拥有的,其中策略A要求被试具有属性A1~A5,策略B要求被试具有属性A3~A7.2类测验情境的设计为:第1类为传统的单策略认知情境,即模拟产生的被试只用策略A的Q矩阵进行答题;第2类为多策略认知情境,即一半被试采用表1中的A策略,另一半被试采用B策略分别进行答题.

3.4 数据模拟

3.4.1 策略的分配 模拟生成1000名被试,采用随机分配方式将2种策略下所有的掌握模式分配给被试.具体而言:在单策略情境下,将所有被试全都平均分配给A策略中的32种掌握模式.在多策略情境下,一半被试分给A策略下的掌握模式,一半被试分配给B策略下的掌握模式,其中不能平均分配的,则随机分派某一掌握模式给被试.

3.4.2 模拟观察反应模式 在实验中,随机失误率(2%、5%、10%)和测验情境(单策略情境、多策略情境)为被试间变量,判别方法(SS-HDD、MS-HDD)为被试内变量,模拟6种条件下的反应数据,每种条件重复30次.以随机失误率2%为例,说明被试观察反应模式的模拟过程:(i)随机生成1000名被试,将被试分配给各测验情境.单策略时,将1000名被试平均分配至A策略上的32种IRP中.多策略时,将1000名被试随机分成2组,各500人,分别随机分配给A策略和B策略;(ii)在A策略和B策略下,分别将500名被试平均分配给本策略的32种IRP;(iii)生成 1000×30 的服从 $U(0,1)$ 的随机数 r :若 r 小于0.01或 r 大于0.99,则将位置所对应的0转换为1,1转换为0,否则不变.

3.5 数据分析

分别用SS-HDD和MS-HDD对单策略测验情境和多策略测验情境的数据进行分析,以考察MS-HDD的稳定性和精确性.

3.6 评价指标

研究采用属性平均判准率(Average Attribute Match Ratio, AAMR)、模式判准率(Pattern Match Ratio, PMR)和解题策略判准率(Strategy Match Ratio, SMR)等作为评价指标.属性平均判准率表示被试在各个属性被正确判为掌握与否的正确比率;模式判准率表示被试属性掌握模式被正确判别的人数比率;解题策略判准率表示正确判别被试使用何种策略的人数比率.公式为 $PMR =$

$$\left(\sum_{i=1}^N N_{i-correct} \right) / N, AAMR = \left(\sum_{i=1}^N \sum_{k=1}^K N_{ik-correct} \right) / (N \times K), SMR = \left(\sum_{i=1}^N N_{i-strategy-correct} \right) / N, \text{其中 } N \text{ 为被试数目, } K \text{ 为属性个数, } N_{i-correct} \text{ 为被试 } i \text{ 是否被归类判别正确, } N_{ik-correct} \text{ 为被试 } i \text{ 在属性 } k \text{ 上被判别归类准确与否, } N_{i-strategy-correct} \text{ 为被试 } i \text{ 的解题策略是否判别正确.}$$

3.7 结果

3.7.1 MS-HDD与SS-HDD之间的比较 表2为使用SS-HDD和MS-HDD分别对各实验条件下重复30次实验的判准率均值.从表2可知,随着随机失误概率的增加,不论哪种情景和条件下,SS-HDD和MS-HDD的判准率都呈下降趋势,这与所有的已有结果都是一致的.因此,不再对失误率与判准率之间的关系进行单独分析,而把重点放在SS-HDD和MS-HDD在各情境下的适宜性分析.

首先,在单策略情境下,从理论角度来说,MS-HDD等同于SS-HDD.结果显示MS-HDD与SS-HDD的判准率差异不大.表2显示,在单策略情境,无论是AAMR、PMR,还是SMR,MS-HDD与SS-HDD的变化趋势均一致.在精确的准确率上来说,SS-HDD与MS-HDD基本无差异,在单策略情境,无论是MS-HDD还是SS-HDD均有较高的判准率,说明MS-HDD在单策略情境下同样适用.

其次,在多策略情境下,SS-HDD的判准率远低于MS-HDD.结合表2和图2的右图,可以明显地看出,在多策略情境,SS-HDD的AAMR、PMR和B策略的SMR均远远低于MS-HDD.具体而言,MS-HDD在策略A和策略B上的判准率均较高,表明MS-HDD能准确地判断被试到底采用了何种策略进行作答,而SS-HDD则无法作此判断,从而把所有的被试都判归A策略.此外,相比MS-HDD,SS-HDD的PMR尤其低,无论在何种条件下,均不足0.50,而MS-HDD的PMR却依然较高,平均达0.8622.由此,在多策略测验情境下,若选用SS-HDD,则会造成对被试AAMR、PMR和SMR的严重误判,应选用MS-HDD为宜.

3.7.2 MS-HDD与MSCD的比较 为使研究结果可比,采用了与涂冬波等^[2]相同的模拟设计.表3是同等条件下GDD和MSCD在单策略和多策略情境下的判别结果.从表2和表3可以看出,MSCD和MS-HDD在各种情景下各判准率的变化趋势基本一致. MSCD对策略A的判准率比MS-HDD稍高,而在B策略的判准率上,两者各有高低,但相差均不大.图3直观地呈现了4种方法在单策略和多策略情境下的PMR变化趋势.由图3可以看出:在单策略情

境下, MS-HDD 和 SS-HDD 的判准率均高于 GDD 和 MSCD; 在多策略情境下, SS-HDD 和 GDD 的判准率均在 0.50 以下, 且差异不大, 而 MS-HDD 与 MSCD 均较理想, 且在失误率 5% 以下, MSCD 要好于 MS-

HDD, 当失误率恰为 5% 时, 两者相等, 但在失误率为 10% 时, MS-HDD 要好于 MSCD. 可见, MS-HDD 与 MSCD 不仅变化趋势一致, 且结果相当, 表明本研究所构建的 MS-HDD 方法具有较好的精确性与适宜性.

表 2 单策略和多策略情境下的 SS-HDD 和 MS-HDD 的判别结果

测验情境	随机失误概率	SS-HDD				MS-HDD			
		AAMR	PMR	A 策略 SMR	B 策略 SMR	AAMR	PMR	A 策略 SMR	B 策略 SMR
单策略	2%	0.993 7	0.979 7	1	/	0.991 5	0.972 5	1	/
	5%	0.985 9	0.927 5	1	/	0.984 5	0.927 0	1	/
	10%	0.970 7	0.861 9	1	/	0.969 9	0.859 9	1	/
	MEAN	0.983 4	0.923 0	1	/	0.981 9	0.919 8	1	/
多策略	2%	0.807 9	0.479 9	1	0	0.972 5	0.923 9	0.963 9	0.971 4
	5%	0.793 6	0.442 4	1	0	0.963 9	0.876 9	0.959 9	0.964 4
	10%	0.772 4	0.424 9	1	0	0.943 9	0.785 9	0.923 9	0.934 3
	MEAN	0.791 3	0.449 1	1	0	0.960 1	0.862 2	0.949 2	0.956 7

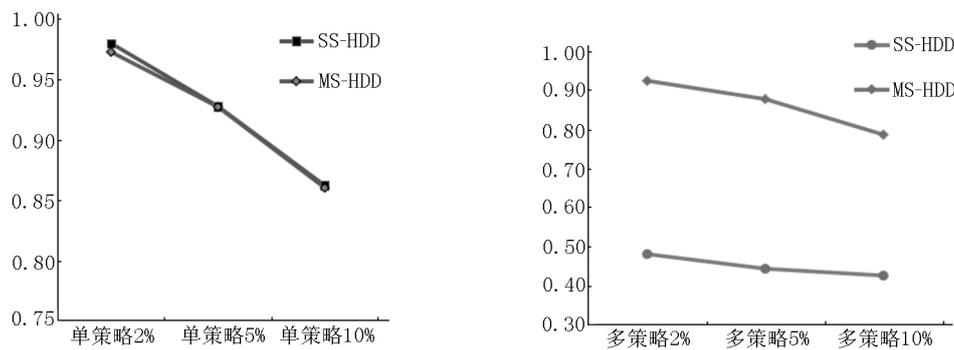


图 2 SS-HDD 和 MS-HDD 在单策略和多策略情境下的 PMR

表 3 单策略和多策略情境下 GDD 和 MSCD 的判别结果

测验情境	随机失误概率	GDD				MSCD			
		AAMR	PMR	A 策略 SMR	B 策略 SMR	AAMR	PMR	A 策略 SMR	B 策略 SMR
单策略	2%	0.989 1	0.956 1	1	/	0.988 4	0.953 2	0.999 4	/
	5%	0.964 5	0.889 5	1	/	0.970 4	0.887 8	0.996 8	/
	10%	0.937 2	0.784 9	1	/	0.936 7	0.780 8	0.992 4	/
	MEAN	0.963 6	0.876 8	1	/	0.965 2	0.873 9	0.996 2	/
多策略	2%	0.714 2	0.506 2	1	0	0.987 2	0.954 4	0.988 8	0.988 4
	5%	0.719 3	0.473 5	1	0	0.962 2	0.876 0	0.959 6	0.963 6
	10%	0.712 6	0.408 5	1	0	0.918 5	0.753 8	0.923 0	0.909 2
	MEAN	0.712 6	0.462 4	1	0	0.955 9	0.861 4	0.957 1	0.953 7

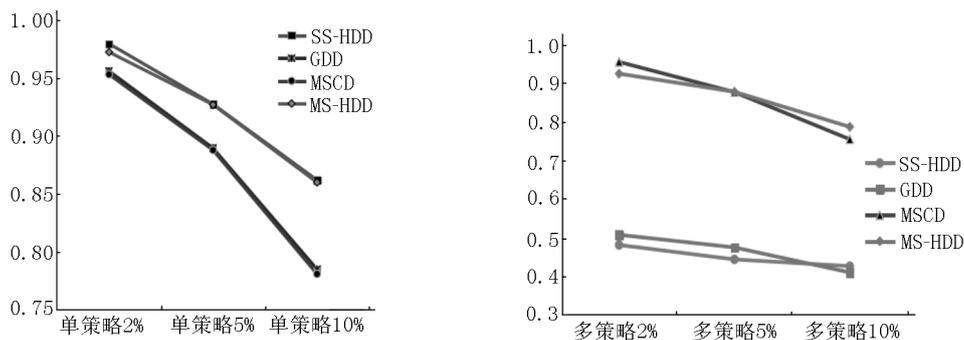


图 3 不同方法在单策略和多策略情境下的 PMR

4 讨论

4.1 MS-HDD 的适宜性

通过借鉴以往多策略研究的思想,将 SS-HDD 这种简便的非参数方法拓展到 MS-HDD,以使其适宜考试形式和评分方式多样化的需求,并通过一个 3 因素的混合设计,以及与 MSCD 的比较,考察了 MS-HDD 的适宜性和稳定性.结果表明:(i) 在单策略测验情境下,MS-HDD 与 SS-HDD 的判别结果几乎等同,这与 GDD 和 MSCD 在单策略情境下的结果几乎重合的结论是一致的;(ii) 在多策略情境下,SS-HDD 无法对与 B 策略有关的掌握模式进行判别,因此其 *PMR* 只能是低于 0.50(因模拟设计中分配了一半的被试使用策略 A).相反,MS-HDD 由于同时考虑了 A、B 这 2 种策略下的掌握模式,因此无论是 *PMR* 还是对 A、B 策略的 *SMR* 的判别都非常高.多策略情境下的结果,与涂冬波^[2]关于 MSCD 的结果趋势也是一致的.由此,MS-HDD 不仅适宜单策略测验情境,尤其适宜多策略情境,且具有较高的判别率.

4.2 MS-HDD 的策略判别率

在多策略测验情境时,MS-HDD 与 MSCD 相比,在失误差率为 2% 的时候,MS-HDD 的 *SMR* 略低于 MSCD,而当失误差率 $\geq 5\%$ 时,MS-HDD 的 *SMR* 与 MSCD 相差不大且略高于 MSCD.出现这一结果的可能原因是 MS-HDD 在进行判别归类时,将与不同 IRP 具有相等距离的 ORP 随机指派给其中的一种 IRP,从而由此导致对策略的误判.罗照盛等^[7]的研究中提出了处理距离相等时的贝叶斯方法,即 B 方法.而本研究只是采用了随机指派的 R 方法.未来研究可考虑采用 B 方法,进一步考察 MS-HDD 策略判别的精确性.

4.3 MS-HDD 的优越性

在单策略情境时,MS-HDD 和 SS-HDD 的 *PMR* 均高于 GDD 和 MSCD,在多策略情境下,MS-HDD 和 MSCD 的 *PMR* 在失误差率不同时,各有高低,但差异不大.这表明,MS-HDD 在各种情景下不仅可以获得与 MSCD 这种参数方法相当的判别率,而且还无需前提假设,无需借助复杂的参数估计算法获得各种结果,只需计算 ORP 与 IRP 之间的距离就可完成

对被试知识状态和策略使用的判别.因此,MS-HDD 不仅适宜多策略测验情境,还具有易于理解和计算简单的优势,更利于 CDA 在实践中的应用.

4.4 未来研究

首先,本研究只应用了 HDD 中的 R 方法进行判别归类,R 方法对具有相同距离的 ORP 进行随机指派,可能会降低判别率,未来研究可以探讨 B 方法的有效性^[7].其次,本研究只考察了 0-1 评分方式的多策略方法,未来可在此基础上拓展到多级评分的多策略方法.最后,能否将两策略时的被试间策略情境扩展至多策略时的混合策略情境,是另一值得研究的问题.

5 结论

本研究将 SS-HDD 拓展至 MS-HDD,并通过多因素的模拟设计考察了 MS-HDD 的精确性与适宜性.结果表明:1) 单策略情境时,MS-HDD 与 SS-HDD 具有等同的判别效果,且判别率均高于相同条件下的 GDD 和 MSCD;2) 多策略情境时,MS-HDD 具有与 MSCD 相当的 *PMR* 和 *SMR*,MS-HDD 适于多策略测验情境;3) 未来研究可进一步将 MS-HDD 拓展至多级评分情境,并考虑对被试混合策略使用的判别.

6 参考文献

- [1] 辛涛,乐美玲,张佳慧.教育测量理论新进展及发展趋势[J].中国考试,2012(5):3-11.
- [2] 涂冬波,蔡艳,戴海琦,等.一种多策略认知诊断方法:MSCD 方法的开发[J].心理学报,2012,44(11):1547-1553.
- [3] Chiu C Y, Douglas J A, Li Xiaodong. Cluster analysis for cognitive diagnosis: theory and applications [J]. Psychometrika, 2009, 74(4): 633-665.
- [4] 祝玉芳,王黎华,丁树良,等.多策略的多级评分认知诊断方法的开发[J].江西师范大学学报:自然科学版,2015,40(4):371-376.
- [5] 康春花,任平,曾平飞.非参数认知诊断方法:多级评分的聚类分析[J].心理学报,2015,47(8):1077-1088.
- [6] Chiu C Y, Douglas J. A nonparametric approach to cognitive diagnosis by proximity to ideal response patterns [J].

- Journal of Classification 2013, 30(2): 225-250.
- [7] 罗照盛,李喻骏,喻晓锋,等.一种基于 Q 矩阵理论朴素的认知诊断方法 [J].心理学报,2015(2):264-272.
- [8] 康春花,任平.聚类诊断分析法诊断正确率的影响因素 [J].中国考试,2015(2):25-32.
- [9] de la Torre, Douglas J A. Model evaluation and multiple strategies in cognitive diagnosis: an analysis of fraction subtraction data [J]. Psychometrika, 2008, 73(4): 595-624.
- [10] 孙佳楠,张淑梅,辛涛,等.基于 Q 矩阵和广义距离的认知诊断方法 [J].心理学报,2011,43(9):1095-1102.
- [11] 李瑜.多选题认知诊断测验编制及多策略的多选题认知诊断模型的开发 [D].南昌:江西师范大学,2014.
- [12] 丁树良,汪文义,杨淑群.认知诊断测验蓝图的设计 [J].心理科学,2011(2):258-265.
- [13] 丁树良,杨淑群,汪文义.可达矩阵在认知诊断测验编制中的重要作用 [J].江西师范大学学报:自然科学版,2010,34(5):490-494.
- [14] Huo Yan, De l T J. Estimating a cognitive diagnostic model for multiple strategies via the EM algorithm [J]. Applied Psychological Measurement 2014, 38(6):464-485.

The Formulation of Multiple-Strategies Hamming Distance Discrimination

LI Yuanbai, ZENG Pingfei, YANG Yakun, KANG Chunhua*

(College of Teacher Education, Zhejiang Normal University, Jinhua Zhejiang 321004, China)

Abstract: The purpose of the paper is to formulate a nonparametric multiple-strategies cognitive diagnosis discrimination method, which called Multiple-Strategies Hamming Distance Discrimination (MS-HDD), to estimate examinees' knowledge states that generated in multiple-strategies of problem solving. The results show that under the single-strategy testing situation, all of the SMR, AAMR and the PMR of MS-HDD method are as similar as SS-HDD's results. Under the multiple-strategies testing situation, the MS-HDD leads to more accuracy in these indexes than the SS-HDD. Under the multiple-strategies testing situation, MS-HDD is compared with MSCD, and the result show that MS-HDD has more better performance than MSCD when the random slip is growing. All in all, the method that extends the SS-HDD method to all for multiple-strategies of problem solving, which is called MS-HDD, was suited to multiple-strategies testing situation and could be accepted in our practical application.

Key words: cognitive diagnosis; multiple-strategies; Hamming Distance Discrimination; PMR

(责任编辑:冉小晓)