

基于前门调整法的“项目式学习”过程 考核设计与课程质量评价

钱 辰¹ 旷 怡¹ 章 兢^{2,3}

(1. 湘潭大学自动化与电子信息学院, 湖南 湘潭 411105;

2. 湖南大学电气与信息工程学院, 长沙 410006;

3. 湘潭大学专业认证指导中心, 湖南 湘潭 411105)

摘要 工程教育是我国高等教育的重要组成部分, 而工程教育专业认证是国际通行的工程教育质量保障制度, 也是实现工程教育国际互认和工程师资格国际互认的重要基础。“项目式学习”是应对工程教育专业认证考察学生能力要求的一种新兴课程教学方法, 因为受学生“惰性”等混杂因子影响, 常规课程质量评价中的考试成绩无法真实反映学生在学习中所能获得的能力提升。本文引入因果推断科学的前门调整法, 在面临难以控制与观测的混杂因子时, 一方面指导优化项目式学习的教学与考核设计, 以改进混杂因子对学习和成果两方面的不利影响; 另一方面用于评价项目教学法对于学生学习成果的真实因果效应。

关键词: 项目式学习; 前门调整法; 因果推断; 专业认证; 课程质量评价

Examination design and course quality evaluation of “project-based learning” process based on front-door adjustment

QIAN Chen¹ KUANG Yi¹ ZHANG Jing^{2,3}

(1. College of Automation and Electronic Information, Xiangtan University, Xiangtan, Hu'nan 411105;

2. College of Electrical and Information Engineering, Hu'nan University, Changsha 410006;

3. Accreditation Guidance Center, Xiangtan University, Xiangtan, Hu'nan 411105)

Abstract Engineering education is an important part of higher education in China, and engineering education professional certification is an internationally accepted engineering education quality assurance system, and it is also an important basis for achieving international mutual recognition of engineering education and engineer qualification. “Project-based learning” is a new course teaching method to meet the requirements of engineering education professional certification to examine students' ability. Because of the influence of students' “inertia” and other hybrid factors, the examination results in the conventional course quality evaluation can not truly reflect the ability improvement that students can obtain in learning. This paper introduces the “front-door adjustment” of causal inference science. When facing the confounding factors that are difficult to control and observe, on the one hand, it guides and optimizes the teaching and examination design of project-based learning to improve the adverse effects of confounding factors on learning and results. On the other hand, it is used to evaluate the real causal effect of project-based teaching method on students' learning results.

Keywords: project-based learning; front-door adjustment; causal inference; professional certification; course quality evaluation

0 引言

我国的高等教育不断发展,自 2016 年 6 月正式加入《华盛顿协议》以来,我国工程教育专业认证工作不断推进,通过认证的专业数量正在不断增长,专业认证成为我国教育领域的一大重点。

现如今很多课程老师为了提高课堂效果,获得更高的课程目标达成度,采用“项目式学习(project-based learning, PBL)”教学方法,以学生为中心,以真实项目或案例为主题,让学生自主学习或小组讨论,从而在学习过程中获得自我能力的提升。但 PBL 教学方法在实际应用中的效果并不理想,因为其中很多同学无法达到自主学习的状态,教学中的作业任务并未自主完成而是借用其他同学的作业去交差,小组讨论中也没有发挥积极作用,但是其在最后考试中反而取得了不错的成绩,所以最后的成绩所反映出来的效果不能真正地表达学生的实际能力。也就是说,有很多难以控制与观测的混杂因子在同时影响着学生的项目式学习过程与最后的考试成绩,项目式学习课程现状因果图如图 1 所示。

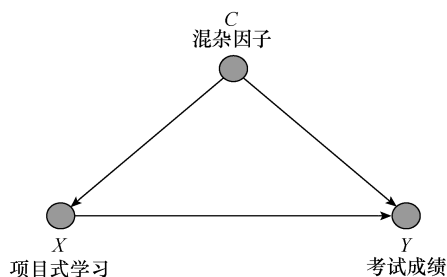


图 1 项目式学习课程现状因果图

由此可见,在面临难以控制与观测的混杂因子时,需要解决两个问题:①是否有清晰易懂的推理模型能够为提出的诸多问题提供明确的答案,设计项目式学习的教学与考核,以改进混杂因子对学习成果两方面的影响;②是否有适应性强的计算公式能够根据所掌握的数据回答因果问题,评价项目教学法对于学生学习成果的真实因果效应。

本文将因果推断科学的前门调整法迁移应用到课程教学设计与质量评价中。一方面,在面临无解的混杂因子时,教师运用因果图,着手寻找不受混杂因子影响的中介变量,也就是新的教学或考核环节,在减少混杂因子影响的同时,能形成使用前门调整公式的条件;另一方面,运用前门调整公式作为教学方法对学生学习成果因果效应的被估量,利用课程观测数据来估计真实的因果效应。

1 理论与方法

1.1 因果理论与因果图

因果理论的发展与因果模型自 1998 年《科学》杂志刊登了一篇名为《大数据的处理程序》最早提出了“大数据(big data)”,短短数十年已发展成为一门重要的学科^[1]。现常用的因果模型通常指由朱迪亚·珀尔等于 2000 年左右提出基本概念,至今已经发展成为一门横跨统计学与计算机学科的重大任务,其重点在于关于干预的建模与推理。Bernhard 将其称作新一代信息革命的燃料^[2]。其主要概念是将事物或者事件用节点进行表示,再用带箭头的线段来表示其因果关系。

图 2 是为了严格处理因果关系问题而针对数据集所提出的一种结构因果模型(structural causal model, SCM)概念的图形化因果模型关联集。其给出的图形定义为,如变量 X 是变量 C 的子变量,即有箭头从 C 指向 X ,则称 C 为 X 的直接因果原因,箭头代表的是 C 到 X 的概率或者说是贡献值,表示有 C 与无 C 时 X 出现的差值^[3]。

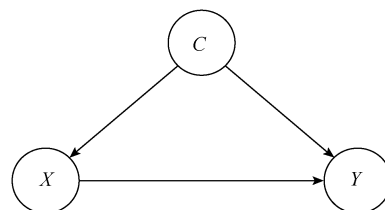


图 2 因果模型

图 2 中,真正的因果效应 $X \rightarrow Y$ 与由叉接合 $X \leftarrow C \rightarrow Y$ 诱导的 X 和 Y 之间的伪相关混合在一起,这里的 C 就叫混杂因子,若 C 作为不可观测的混杂因子,影响 $X \rightarrow Y$ 的因果效应时,因果效应就无法通过条件期望计算,这时需要采用随机试验计算因果效应或者采用传统统计学方法进行回归分析。但在工程教育领域中,采用随机试验分析课程目标达成度的方法无法使用,工程教育具有全员覆盖性,随机试验不现实,同样用回归分析的方法计算因果效应的方法过于繁琐,不便于在教育领域的推广与应用。因果图提供了一种将数据与因果信息相结合的可选语言,可使用后门调整法、前门调整法等直观的方法解决因果推断问题,这使因果图非常适合在工程教育领域应用。

1.2 do 算子与后门调整

因果学可以将认知能力分为三种层级:观察能力(seeing)、行动能力(doing)和想象能力

(imagining)^[4], 其中第二层级的行动能力代表实施“干预”。“干预”就是控制一个变量, 以影响其他变量, 是判定因果关系的一个重要操作。如果控制一个变量, 导致其他变量的概率分布, 则可以说明前者是后者的因^[5]。通过干预可以确定多个变量间因果关系的存在性, 这也是现如今很多科学方法常用的手段, 例如控制变量法及随机对照实验等。然而在很多情境下, 实施干预去确定因果关系的成本及风险都难以估计, 如前文所言, 在具有全员覆盖性的教育领域, 实施干预的举措就难以完成。这就使得在这些情境下, 需要通过一些举措观测现有数据去预测干预的效果。

这里需要调用“do 算子”来表达预测实施了干预的结果, do 算子的引用是因果理论与传统统计学最为显著的区别。“do”代表了行动, 可以表达因果关系第二层级的一个经典问题, 即“如果我们实施了某种行动, 这将会带来何种结果?”, 缺乏 do 算子的条件概率仅仅反映了因果关系第一层级的观察所得到的结果。其中 do(X)对于因果图来说就是删除指向 X 的所有箭头。

例如图 2 中, 当混杂因子 C 可测量时, 可以使用后门调整法计算 X 对 Y 的因果效应 $P(Y|do(X))$ 。

后门路径可以定义为所有 X 和 Y 之间以指向 X 的箭头为开始的路径, 如果可以阻断所有的后门路径, 也就意味着完成了去混杂^[6]。控制 C 即可阻断后门路径, 后门调整就是通过对去混变量集 C 的控制阻断从 X 到 Y 的所有后门路径; 对于数学公式, 此时, do(X)等同于 see(X), $P(Y|do(X), C) = P(Y|X, C)$, 表达的意思是, 在控制了一个充分的去混因子集 C 之后, 留下的相关性, 就是真正的因果效应。

由 do 算子运算可以推导出后门调整公式为

$$\begin{aligned} P(Y|do(X)) &= \sum_C P(Y|do(X), C) \cdot P(C|do(X)) \\ &= \sum_C P(Y|X, C) \cdot P(C|do(X)) \\ &= \sum_C P(Y|X, C) \cdot P(C) \end{aligned} \quad (1)$$

此公式计算的是变量 X 对于变量 Y 的因果效应, 其中变量 C 为变量 X 对变量 Y 因果效应中的混杂因子。通过公式, 可以在 C 可观测的情况下, 直接通过数学计算得到 X 到 Y 的因果效应。

1.3 前门调整法

在 C 不可测量的情况下, 已无法通过后门路径简化计算, 此时可以选用前门调整法, 在 C 不可测

量的情况下去除混杂, 从而计算 X 到 Y 的因果效应。

如果一个变量集合 M 满足以下条件: M 切断了所有 X 到 Y 的有向路径; X 到 M 没有后门路径; 所有 M 到 Y 的后门路径都被 X 阻断, 则称变量集合 M 满足有序变量 (X, Y) 的前门准则。也就是说, 当计算因果效应时发现 X 对 Y 的因果效应被一组变量 C 混杂, 又被另一组变量 M 介导, 且变量 M 不受变量 C 的影响, 便能知道这里可以利用前门调整法从观测数据中估计 X 对 Y 的因果效应。当意识到这点之后, 在面临无解的混杂因子时, 首先应想到寻找不受混杂因子影响的中介变量^[7]。

加入中介变量的因果模型如图 3 所示, 此模型满足前门调整法的使用条件, 所以 X 到 Y 的因果效应 $P(Y|do(X))$ 可以使用前门调整公式表达为

$$P(Y|do(X)) = \sum_m P(M=m|X) \sum_{x'} P(M=m|X=x') P(X=x') \quad (2)$$

此公式计算的是变量 X 对于变量 Y 的因果效应, 其中变量 M 是变量 X 对变量 Y 因果效应的中介变量。

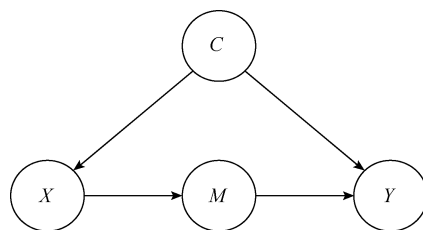


图3 加入中介变量的因果模型

2 因果模型搭建与分析

2.1 因果模型搭建

回到图 1 的因果图所展示出的问题, 因为混杂原因无法真实找到项目教学法对于学生学习效果的因果效应, 所以本文目的即是去除影响两个变量的混杂因子。现如今的评价方式大多使用考试成绩作为反映学生能力的具体体现, 将考试成绩作为课程目标达成情况的考察标准, 但是考试成绩可由多种途径提高, 例如老师考前划重点、学生考前刷真题等, 这类因素可以使一些实际能力未能达标的同学由考试成绩的数值反映为达标。

若想要知道项目式学习本身对学生学习效果的因果效应, 则需主要研究 $P(\text{考试成绩}|do(\text{项目式学习}))$, 现在教育常规评价的是 $P(\text{考试成绩}|\text{项目式学习})$, 这种条件概率公式表示学习产出达成的概率

(P) 是以观察到学生参加了实验为条件的。注意 $P(\text{考试成绩}|\text{项目式学习})$ 与 $P(\text{考试成绩}|\text{do}(\text{项目式学习}))$ 完全不同, 观察到 (seeing) 和进行干预 (doing) 有本质的区别, 两者的混淆成为高等教育评价社会失信之源。完全由 $P(\text{考试成绩}|\text{项目式学习})$ 统治的高等教育评价是荒诞的。

因此, 本文使用因果推断的方法计算课程目标达成度, 即计算项目式学习对于考试成绩的因果效应, 需要排除影响项目式学习与考试成绩间因果效应的混杂因子。现将“学生作业抄袭”“学生在小组讨论中划水”“学生考前刷真题”“作弊”等因素归类为“惰性”, 作为项目式学习对于考试成绩因果效应中的混杂因子。然而“惰性”大多为学生自身主观的消极因素, 无法观测, 无法使用后门调整法去除混杂。由朱迪亚·珀尔的前门调整法可知, 在面临无解的混杂因子时, 应立即着手寻找不受混杂因子影响的中介变量, 从而使用前门调整法从观测数据中估计 X 对 Y 的影响^[8]。

对于项目式学习与课程目标达成情况的中介变量, 本文选用“成果展示”。“成果展示”指的是学生在完成自己以项目为基础的作业后, 在课堂上进行展示并由其他学生进行投票打分。以“成果展示”作为项目式学习到课程目标达成情况的中介变量, 是因为现在大多数的老师意识到以面向成果导向的工程教育专业认证注重的是以解决实际问题为目标的学生能力, 而不是传统的以考察学生知识点掌握能力为终极目标的分数教育, 由此, 需要一个过程性评价^[9]。所以这里要求学生在课堂上展示自己所完成的以项目为基础的作业。

某校信息安全课程因果模型如图 4 所示, 在这里, X 代表项目式学习, Y 代表课程目标达成情况, C 代表混杂因子惰性, M 代表成果展示。

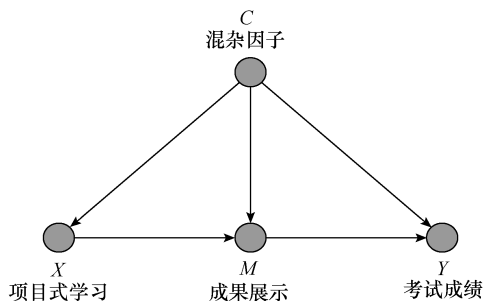


图 4 某校信息安全课程因果模型

2.2 基于前门调整法的模型分析

观察图 4 可以知道, 因为未采取任何措施, 惰性同样影响着学生的成果展示成绩, 即混杂因子 C

影响了中介变量 M , 所以 C 有指向 M 的箭头, 这导致混杂因子无法在引入中介变量后得到有效的控制和去除。

所以如果要使用前门调整法, 就必须屏蔽混杂因子对于成果展示的影响, 取消掉 C 到 M 的箭头。课程可以在大纲中明确要求每人至少参加一次以项目为基础的成果展示, 将不参加成果展示则不能参与期末考试作为硬性要求, 因为项目式学习作业次数多, 且难以每次严格管理, 而成果展示每个人只有一次, 可以做到严格管理成果展示这一环节且精确评分, 这些举措就阻断了惰性这一混杂因子对于过程评价即“成果展示”的影响, 消除了因果图中 C 指向 M 的箭头。改进后的课程因果模型如图 5 所示, 这一举措使中介变量“屏蔽”了混杂因子的影响, 因此适合使用前门标准来分析项目式学习对于考试成绩的因果效应。

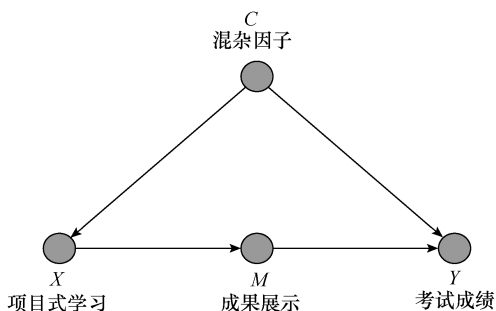


图 5 改进后的课程因果模型

对于这样的因果模型, 可以发现公式中看不到 C (惰性因子) 的存在, 这使此方法在未采集到任何数据的时候就成功地排除了混杂因子 C 。其次, 因果推断引擎作为一种问题处理机器, 以“假设”为基础, 接收“问题”输入, 产生输出“被估量”^[10]。被估量可视作一种针对问题中目标量的计算方法。前门调整和后门调整公式就是两个被估量。公式左边代表提出问题“ X 对 Y 的影响是什么”, 而公式右边是被估量, 即是回答左边提出问题的一种方法, 此被估量以条件概率的形式表示, 不包含关于 do 算子的有关数据, 也就是说与关于实际干预的数据无关, 只包含观测到的数据^[11]。这意味着它无需依赖随机对照实验, 可以直接根据数据估计出来。

3 数据分析

3.1 数据来源

本文选用某校 2019 级通信工程班信息安全课程 101 人成绩, 成绩包括项目式学习的作业成绩,

学生以项目为基础的成果展示的成绩，以及考试成绩。对于 X ：“0”“1”分别代表着学生项目式学习平时成绩不合格和平时成绩合格。对于 M ：“0”“1”“2”分别代表着学生在成果展示的成绩中取得“不及格”“及格”与“优秀”。对于 Y ：“0”“1”“2”分别代表着学生在期末考试中取得“不及格”“及格”与“优秀”。

对于此课程的课程目标达成度分析，本文采用将 $P(Y \geq 1 | do(X=1))$ 的值大于 70% 视为课程目标达成度合格的方法。

3.2 数据处理

将 101 位同学的各项成绩进行分类，在平时成绩合格和不合格的情况下，成果展示成绩取得“不及格”“及格”与“优秀”人数见表 1。

表 1 各平时成绩级别下的成果展示成绩人数

单位：人

X	$M=0$	$M=1$	$M=2$	总
0	15	10	1	26
1	10	43	22	75

由表 1 可以计算，在项目式学习平时成绩不合格的同学中，成果展示环节成绩不及格的概率为 $P(M=0|X=0)=57.69\%$ ，成绩及格的概率为 $P(M=1|X=0)=38.46\%$ ，成绩优秀的概率为 $P(M=2|X=0)=3.85\%$ ；在项目式学习平时成绩合格的同学中，成果展示环节成绩不及格的概率为 $P(M=0|X=1)=13.33\%$ ，成绩及格的概率为 $P(M=1|X=1)=57.33\%$ ，成绩优秀的概率为 $P(M=2|X=1)=29.33\%$ 。

对已有数据进行进一步处理，得到在项目式学习平时成绩两个不同成绩级别及成果展示环节三个不同成绩级别的条件下，考试成绩取得“不及格”“及格”与“优秀”人数见表 2。

表 2 平时成绩和成果展示各成绩级别下的不同考试成绩级别人数

单位：人

X	M	$Y=0$	$Y=1$	$Y=2$	总
0	0	15	0	0	15
0	1	4	6	0	10
0	2	0	1	0	1
1	0	6	4	0	10
1	1	1	38	4	43
1	2	0	7	15	22

将表 2 中数据代入后门调整公式可知，如果学生按照要求分别取得成果展示成绩为“不及格”“及格”与“优秀”，那么他们在这种条件下，可能取得不同的考试成绩级别的概率见表 3。

表 3 实验假设后各成果展示成绩级别下的考试成绩级别概率

单位：%

$P(Y do(M))$	$Y=0$	$Y=1$	$Y=2$
$M=0$	70.29	29.70	0
$M=1$	12.02	81.07	6.91
$M=2$	0	49.37	50.63

最后由前门调整公式可知，如果学生按照要求参加项目式学习并且平时成绩合格的话，那么学生在考试中取得成绩的概率见表 4。

表 4 实验假设后项目式学习不同成绩级别下的考试成绩级别概率

单位：%

$P(Y do(X))$	$X=0$	$X=1$
$Y=0$	36.70	13.52
$Y=1$	50.21	64.92
$Y=2$	4.61	18.81

在这里，数据中 $P(Y|do(X=0))$ 可以表达为“如果学生未能按照要求完成项目式学习任务的话，那么其在考试成绩中所能取得某种成绩的概率为……”，同理， $P(Y|do(X=1))$ 可以表达为“如果学生按照要求完成项目式学习任务的话，那么其在考试成绩中所能取得某种成绩的概率为……”。

从以上数据可以看出，学生如果按照要求完成项目式学习任务的话，那么在期末考试中及格的概率为 $P(Y=1|do(X=1))=64.92\%$ ，其在期末考试中获得优秀的概率为 $P(Y=2|do(X=1))=18.81\%$ 。所以 $P(Y \geq 1 | do(X=1))=83.73\%$ 。

现如今的国内课程目标达成度分析普遍使用平均分算法、15%分位法等，类似方法太过于考虑数据而忽略了数据生成过程中的混杂因素。

去除与未去除混杂的课程目标达成度情况见表 5，对比不使用因果推断仅用条件概率算出的课程目标达成度情况 $P(Y=1|X=1)=65.33\%$ ， $P(Y=2|X=1)=25.33\%$ ，所得出的 $P(Y \geq 1|X=1)=90.66\%$ ，有明显的区别。若按条件概率所得到的课程目标达成度计算，混杂原因会使课程目标达成度的值普遍较高，这可

能让一些实际并未合格的课程在未去除混杂的情况下也在条件概率的数据下显示合格,这明显不符合工程教育专业认证的目的,因此去除混杂计算课程目标达成度的因果推断方法更符合工程教育专业认证的要求。

表 5 去除与未去除混杂的课程目标达成度情况

类别	$P(Y \text{do}(X=1))/\%$	$P(Y X=1)/\%$
$Y=1$	64.92	65.33
$Y=2$	18.81	25.33
总	83.73	90.66

由以上数据可得, $P(Y \geq 1|\text{do}(X=1))=83.73\% > 70\%$,所以可以知道本课程的课程目标达成度合格。去除混杂后的数据直接反映了学生参加课程教学对于期末考试成绩的因果效应,而去除“惰性”混杂因子后的考试成绩可以反映出学生自身的学习成果产出,也就说明了此数据更能反映学生的课程目标达成度情况。

3.3 变量选取分析

本文中变量 X 本质为项目式学习这种教学方法,而度量一种教学方法主要是观察学生自身的改变,通常可以选用参与度来度量教学方法。参与度的高低需要数值来进行区分,故本文选取平时成绩的合格与否,赋值“1”“0”作为参与度高或低的区分。实际上依据课程的教学情况,可以对参与度有更多的档次分层。

对于中介变量 M ,本文选用以项目为基础的成果展示,在项目式学习中起到两大关键作用:首先是对干预有预测作用,即在不进行随机试验的情况下也可以模拟随机试验的效果;第二,在教学大纲中加入成果展示环节,对学生在学习的投入有积极的促进作用。在不同的课程中可以根据实际教学情况有不同的选取方式,例如在需要进行实验的课程中选取实验作为中介变量,但要保证混杂因子对中介变量无影响,需对实验的环境与评价方式进行一些设计,比如采用人工智能评价方法去评测结果。

在本文中,变量 Y 是采用学生的期末考试数据,但实际上项目式学习这种教学方式本身就产生了多元化的过程和考核结果,在不同课程的课程目标达成情况分析中,可以选取不同的考核数据,或是多次的考核结果数据,因为在本文图 1 中,变量 Y 的选取并不影响变量 X 是变量 Y 的因,变量 Y 是变量 X 的果,同理,在图 5 中,变量 Y 既是变量 X 的果,

也是变量 M 的果,变量间的因果关系都不会因为 Y 的构成而发生变化。

4 结论

本文引用了因果推断方法中的 do 算子及后门标准、前门标准的概念,将之迁移到教育领域,提出构建基于前门调整法的“项目式学习”过程考核设计与课程质量评价方法,在工程教育具有全员覆盖性、无法实施随机对照试验的前提下,采用基于观测数据的自然实验方法,有效去除数据关联中的混杂影响,从而推断出项目式学习对学生学习成果的真实因果效应。本文方法是结合教学实践的有益探索,对专业认证实施有积极促进作用。

参考文献

[1] 王东明,陈都鑫. 因果推断: 起源和发展[J]. 控制工程, 2022, 29(3): 464-473.

[2] 王天思. 大数据中的因果关系及其哲学内涵[J]. 中国社会科学, 2016(5): 22-42, 204-205.

[3] 白永梅,孙华鸽,杜建. 知识图谱: 一种系统性构建因果图的方法[J]. 首都医科大学学报, 2022, 43(4): 584-591.

[4] 朱迪亚·珀尔,达纳·麦肯齐. 为什么[M]. 江生,于华,译. 北京: 中信出版社, 2019.

[5] SCRIVEN M. “The methodology of evaluation”. In perspectives of curriculum evaluation[M]. Chicago: Rand McNally and Company, 1967.

[6] 金华,方积乾. 因果推断中的混杂控制[J]. 生物数学学报, 2001(3): 362-366.

[7] PEARL J. Causality: models, reasoning, and inference[M]. Camb: Cambridge University Press, 2000.

[8] GILLIES D. Causality: models, reasoning, and inference by Judea Pearl[J]. The British Journal for the Philosophy of Science, 2001, 52(3): 613-622.

[9] 郭娇,吴寒天. 大数据时代的因果推断——教育政策评估的新路径[J]. 重庆高教研究, 2022, 10(4): 39-48.

[10] 涂良川. 因果推断证成强人工智能的哲学叙事[J]. 哲学研究, 2020(12): 110-121.

[11] 李家宁,熊睿彬,兰艳艳,等. 因果机器学习的前沿进展综述[J/OL]. 计算机研究与发展, 2021.

收稿日期: 2022-07-01

修回日期: 2022-09-02

作者简介

钱 辰(1998—),男,安徽省黄山市人,硕士研究生,主要研究方向为人工智能。