

统计学视角的学生课下学习模式及干预模式探究

韩锋¹, 易丹辉^{2*}, 张景胜^{1*}, 张琳琳¹, 李青艳³, 高洪力¹

1. 北京联合大学, 北京 100101

2. 中国人民大学, 北京 100872

3. 北京市陈经纶中学保利分校, 北京 100024

DOI:10.61369/EDTR.2025050027

摘 要 : 高速发展的互联网给我们带来了便利, 也带来一些挑战。混合式教学、项目式教学等模式下, 如何充分利用好课余时间, 在“双减”政策下, 既能减去不必要的负担又能高效地获取知识对学生的发展尤为重要。当今家长们面临着工作压力及生活压力, 同样需要减负。如何借助学生之间的互助互律恰当发挥小组学习的功效非常重要, 而学习小组的恰当选择无疑需要借助科技的力量。本文借助社交网络方法与因果推断方法相融合, 给出学生课下学习模式的一些建议, 从而, 给实际的精准教与学提供借鉴和参考。

关 键 词 : 因果推断; 社交网络; 学生课下学习模式; 干预模式; 度中心性算法

An Exploration of Students' After-class Learning Patterns and Intervention Models from a Statistical Perspective

Han Feng¹, Yi Danhui^{2*}, Zhang Jingsheng^{1*}, Zhang Linlin¹, Li Qingyan³, Gao Hongli¹

1. Beijing Union University, Beijing 100101

2. Renmin University of China, Beijing 100872

3. Beijing Chenjinglun High School. Baoli, Beijing 100024

Abstract : The rapid development of the Internet has brought us both convenience and challenges. It is particularly important for students' development to make full use of off class time in blended learning, project-based teaching and other modes, and to efficiently acquire knowledge while reducing unnecessary burdens under the "double reduction" policy. Today's parents are facing work and life pressures, and they also need to reduce their burden. It is very important to leverage the mutual assistance and discipline among students to effectively utilize group learning, and the appropriate selection of learning groups undoubtedly requires the power of technology. This article combines social networking methods with causal inference methods to provide some suggestions for students' extracurricular learning patterns, thus providing reference and guidance for practical precision teaching and learning.

Keywords : causal inference; social networks; student extracurricular learning patterns; intervention patterns; degree centrality algorithm

引言

2021年7月, 中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》^[1]并提出双减: 即减轻义务教育阶段学生作业负担、减轻校外培训负担。“双减”政策实施之后, 如何有效的利用课后时间是每个家庭和教师及同学颇为关心的议题。学生课下有效学习模式的探索无疑给学生的学业提供保障, 同时, 也有助于发挥“双减”政策的真正效应。

数字化时代的教学活动中积累了大量观察性数据, 为我们更好的教与学提供了参考, 但因为数据量日益增大, 如何恰当的梳理数据无疑需要借助统计学方法。前期我们与某些中小学一线老师及高中大学老师交流, 整理出学生课下学习以期末复习为例, 常有几种方式小组讨论共同复习、学生自主复习、家长或老师课下督促及答疑复习等等。教师及家长常常关心各种方式的有效性, 以某个科目为例, 大家常关心不同课下学习方式的复习效果即学生成绩的差异, 也即因果效应。

基金项目: 本论文受国家社科基金后期资助项目(18FTJ003)资助; 北京联合大学教改项目《大学数学教学中项目式学习的探究》研究成果(JJ2025Z008)。

作者简介: 韩锋(1980.04-), 女, 黑龙江省肇东市人, 博士, 副教授, 主要的研究方向为因果推断、复杂数据的统计方法研究等, 邮箱: hanfeng1661@163.com。

通讯作者:

易丹辉, 女, 教授, 博士生导师, 主要研究方向: 风险管理与保险、预测与决策, 邮箱: xueyi905@aliyun.com

张景胜(1971.09-), 男, 甘肃会宁县人, 硕士, 副教授, 主要研究方向为数学建模, 邮箱: E-mail: zjs918@126.com

要想比较两组或多组的某种课下复习方式的效应，最简单的是进行随机化试验^[2]（Fisher，1925）。但随机化试验实际中未必可行，人是有行为能力的主体，在实际强行指定学生课下复习方式有时候可能会适得其反。

本文基于的是反事实模型^[3]（Rubin，1974），解决了现实不能进行随机化试验的情况。同时，应用倾向评分逆加权方法^[4]（Hirano，Imbens，& Ridder，2003）进行了比较组效果对比，从而，从宏观上探寻适宜的学生课下学习模式。

为了考虑到个体差异及更精准的进行学生课下学习模式的探索。我们借助社交网络方法^[5]（Ding 等，2019），寻找社交网络意见领袖^[6]（郭奕等，2021），从而有的放矢的恰当引导学生们的课下学习。

一、学生课下学习模式效应评价及估计方法

（一）分析的数据采集方法

基于班级现有的数据即可，也称观察性数据。如某班某门课程的成绩作为关注的结局变量；学生课下学习模式如有小组讨论共同学习、教师及家长课下督促引导学习、学生自己独立学习这三种情况，学生课下学习模式作为处理变量，这里取值有三个，在实际的比较中我们可以两两比较；协变量的选取需要一线教师根据实际情况来具体定，如学生的年龄、父母的文化程度、父母的职业、学生的月考成绩、期中考试成绩、平时学校学习状况等。数据整理成如下结构化数据集。（表一）

表一 学生课下学习模式分析数据集

ID	学生的年龄	父母的文化程度	父母的职业	月考成绩	期中考试成绩	平时学校学习状况	学生课下学习模式（处理变量）	期末考试成绩（结局变量）

在实际中通常需要隐去学生的真实姓名和学号信息，视实际情况增加分析的协变量，并量化各个协变量。这里存在一个问题是如何均衡两比较组的协变量，大家很清楚观察性研究两比较组的协变量（混杂因素）通常不能认为同分布的，这里我们介绍一个在因果推断领域均衡混杂因素的常用方法即倾向评分。

（二）比较组混杂因素均衡方法

倾向评分^[7]（Propensity Score）常常用来均衡观察性研究两比较组的混杂因素。其定义如下：

研究对象 i 在混杂因素 X_i 条件下被分配到处理组的条件概率，具体如下：

$$p(x_i) = P(T=1 | X_i = x_i).$$

这里需要注意的是，因为我们后续应用的是倾向评分逆加权方法进行效应估计，因而，倾向评分的取值我们要求位于（0，1）区间。

倾向评分的具体计算，最简单可以应用 logistic 回归：

$$\log \frac{P(T=1|X)}{1-P(T=1|X)} = \alpha_0 + \alpha_1 T + \alpha_2^T X.$$

如下表计算出各样本的倾向评分的值，要均衡各比较组的混杂因素，只需要均衡各个比较组的倾向评分。

表二 带有个体倾向评分的数据集

ID	学生的年龄	父母的文化程度	父母的职业	月考成绩	期中考试成绩	平时学校学习状况	倾向评分值	学生课下学习模式（处理变量）	期末考试成绩（结局变量）

（三）学生课下学习模式效应的估计

这一部分我们来估计学生课下学习模式的效应，也即学生课下学习模式这个处理变量的因果效应。前述我们借助倾向评分（Propensity Score）方法均衡了各比较组的混杂因素，接下来应用加权倾向评分来估计处理的平均因果效应^[4]（Hirano，Imbens & Ridder，2003）。

由加权倾向评分的基本思想，Hirano，Imbens & Ridder（2003）得到加权后平均因果效应或平均处理效应为，也即我们估计的学生课下学习模式的平均效应：

$$ATE = E \left[\frac{TY}{p(X)} - \frac{(1-T)Y}{1-p(X)} \right].$$

针对一个具体样本，加权后平均处理效应的估计：

$$\hat{ATE} = \frac{1}{N} \sum_{i=1}^N \left(\frac{T_i Y_i}{p(X_i)} - \frac{(1-T_i) Y_i}{1-p(X_i)} \right).$$

两比较组学生课下学习模式平均效应的估计出来后，根据取值与0的关系来判别不同学习模式的效应，为了更精确度量这种效应，可以应用 bootstrap^[8]（Efron 等，1993）方法估计方差。估计结果可列为如下表三。

表三 学生课下学习模式平均效应的估计

效应（学习小组与自主学习效应比较）	估计值	置信下限	置信上限	Pr > t

在一定的显著性水平或置信水平下，考察不同学习模式比较效应，如果效应估计值显著大于0且置信区间不覆盖0且均大于0，如果处理组为学习小组方式课下学习且控制组为学生自主学习，则表明在一定的置信水平下学习小组学习方式优于学生自主学习，其他课下学习方式比较类同。从而，选择一个较优的课下学习方式。

二、学生课下学习模式效应干预方法研究

正如教学有法而无定法，即使我们根据实际情况获取了较优

的课下学习方式，同时，教师和家长也有必要进行有效的引导和干预，从而，让课下学习更有效、更有助于学生的正向成长。因而我们有必要了解学生的社交网络，特别是在社交网络中对同学观点或行为具有极强的引导力和影响力的学生，即前述提到的社交网络意见领袖。

社交网络（Social Network）是指由许多节点组成的社交结构，包含个体间因互动而形成的相对稳定的关系体系（Ding 等，2019）。其可以直观展示同学们课下学习之间的网络关系，可能有多中心节点，也就有多小组。我们关心社交网络意见领袖可以借助节点的中心性来判断。我们介绍一种网络节点排序算法，度中心性算法（Ding 等，2019）。

度中心性算法是指：

我们假设社交网络为 $G=(E,L,\omega)$ ，其中 E 为网络中的节点集表示为 $E=\{e_1,e_2,\dots,e_n\}$ ；

L 表示网络中的边集可表示为 $L=\{(e_i,e_j)|e_i,e_j\in E;i\neq j\}$ ； ω 表示边 (e_i,e_j) 的权重，记为 $\omega=\{\omega(e_i,e_j)|e_i,e_j\in E;i\neq j\}$ 。

节点 e_i 的度中心性为： $CD(e_i)=CD^{\text{in}}(e_i)+CD^{\text{out}}(e_i)$

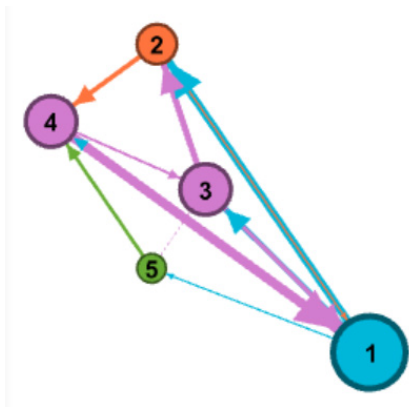
$CD^{\text{in}}(e_i)$ 表明节点的入度中心性， $CD^{\text{out}}(e_i)$ 表明节点的出度中心性。

以五个同学为例，不妨同学间课下交流情况，以某门课程期末复习前重点题学习为例，同学甲督促或给乙答疑过1次记为1，答疑过10次记为10，具体数据梳理如下表四。

表四 学生课下学习交流情况数据集

ID	1	2	3	4	5
1	0	30	23	15	10
2	10	0	0	20	0
3	20	25	0	0	0
4	35	0	13	0	0
5	0	0	0	16	0

我们以一个班40人为例，可以想象这个表格看起来很不方便，实际中应用社交网络方法可以可视化展示，便于一线教学老师及时识别网络情况。以表四为例，具体的社交网络图可表示如下图一。



图一 .课下讨论的社交网络

由图一可以看出，ID为1的同学课下交流活跃度最高，并且由边的方向可以看出，对其余4名同学影响较大；3和4的整体活跃度接近，但4仅对1有较大影响，对3有些影响，3对1、2有影响。

这样我们借助中心节点找到学生课下学习交流中的学生集群及每个集群中的意见领袖。从而，我们借助这些意见领袖及时交流，了解全班动态及大家课下整体情况，进行有效干预。

三、结束语

本文针对教师、家长及社会关注的学生课余时间学习的效应问题进行研究，借助现代统计学方法来对学生课下学习模式及干预模式进行探究。具体应用实际教学中积累的触手可及的观察性教学数据进行分析，借助倾向评分逆概率加权方法估计平均效应，同时给出学习模式比较后的评价方法。针对大家广泛关注的学生社交网络群体情况，为了使教师、家长及社会更有效对学生课下学习进行有效干预，借助社交网络方法直观展示了班级群体的整体情况，实际中可能形成几个交往密切的网络图，每个网络中借助中心性了解每个社交网络中的意见领袖，使沟通更高效、更顺畅。

参考文献

[1]中共中央办公厅 国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》. 央视新闻 [引用日期2021-07-24].
[2]Fisher, R.A. Statistical Methods for Research Workers[M].Oliver & Boyd,1925.
[3]Rubin, D.B.Estimating causal effects of treatments in randomized and nonrandomized studies[J].Journal of educational Psychology, 1974,66:688-701.
[4]Hirano, K., Imbens, G. and Ridder, G.Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score[J].Econometrica, 2003,71(4): 1161-1189.
[5]Ding R X, Wang X, Shang K, et al. Social network analysis- based conflict relationship investigation and conflict degree-based consensus reaching process for large scale decision making using sparse representation[J]. Information Fusion, 2019,50:251-272.
[6]郭奕, 徐亮, 熊雪军. 社交网络中意见领袖挖掘方法综述 [J]. 计算机科学与探索, 2021, 15 (11): 2077-2092.
[7]Rosenbaum, P. R., and Rubin, D. B.The Central Role of the Propensity Score in Observational Studies for Causal Effects[J].Biometrika, 1983,70:41-55.
[8]Efron, B., Tibshirani, R. (1993), An introduction to the Bootstrap. London: UK: Champman & Hall.