

# 基于Markov链的在线课程教学评价成效分析

沈 侯 王立彬

(南京林业大学 江苏南京 210037)

**摘要:** 本文通过对师生发布线上调查问卷、学生评价、学生成绩等方面, 汇总教师的各项教学评价指标数据, 并统一化处理, 通过建立Markov链模型, 排除学生自身学习基础的差异性、不同教师出题的难易程度不同等因素对教学评价结果的影响, 从而较为客观地对教师线上课程的教学效果进行成效分析和评价。

**关键词:** 马尔科夫链 在线课程 教学评价

**中图分类号:** G434 **文献标识码:** A

**DOI:** 10.12218/j.issn.2095-4743.2022.29.056

## 引言

教学评价是研究教师教学效果和学生学习效果的价值过程, 也是一个影响因素多、因变量多的非线性过程, 起到了定向和调控、诊断和预测的功能。传统课题的教学评价基于传统的教学模式, 根据学生考试期中或期末成绩来评价教师的教学效果<sup>[1]</sup>。从2012年世界高等教育开启“慕课元年”伊始, 课题教学的变革也带动着教学评价的变革。尤其2020年全球爆发的新型冠状病毒疫情, 基于“停课不停教、停课不停学”的原则, 将在线课程推向了发展和普及的高潮。

如今随着互联网和移动学习的迅速发展, 越来越多的教学者、学习者开始选择在线教学平台进行课程学习, 如国内的中国大学慕课、网易云课堂, 国外的Coursera等。远程教育的优点是随时随地的教学, 但是缺点则在于教学者很难直观地观察到学习者的学习状态, 且由于学生群体规模庞大, 与教师素不相识, 教学者课后也很难采集评价反馈数据。线上课程教学过程中, 学生无法和教师面对面授课、教师无法实物教学、授课环境复杂等因素让在线教学有了传统课堂无法预测到的特殊性。而疫情过后, 线上线下混合式教学已然成为今后课堂教学的发展趋势, 各自的研究中尝试构建了不同的在线课程质量评价体系, 但在研究方式上总体呈现“自上而下”主观评价的研究特点, 多以对学习者的问卷调查/访谈、邀请专家打分、分析国内/国外教育部发布的相关文件为主。因此, 建立科学合理的评价体系, 成为当下教育教学工作者亟需解决的任务。本文将Markov模型应用到在线课程教学评价中, 采取学生成绩“标准化”的方法排除学生基础的差异以及试卷难易程度不同对评价结果的影响,

客观地评价教师的教学质量, 从而指导老师面临新型的在线课程教学, 应该如何调整自己的上课方式、授课内容以及教学行为<sup>[2]</sup>。

## 一、在线课程发展历程和评价现状

2001年, 美国麻省理工学院启动了开放式课件项目, 自此, 开放教育资源正式开启并日益增长。2012年, 出现了edX、Coursera和Udacity三大社会公开开放的在线课程平台, 极大地推动了MOOC的发展。MOOC的出现, 像一个扳手一样撬动着大学校园的围墙, 为老师和学生们提供了全新的知识获取方式。与此同时, 哈佛大学、加州大学等高校也在尝试一种更完善的课程类型——SPOC (Small Private Online Course 小规模限制性在线课程)。

为了顺应时代发展的需要, 近年来, 我国一批高水平大学也开始进行大规模在线开发课程的建设与应用, 如上海交通大学的“好大学在线”、清华大学的“学堂在线”等等。这些平台汇集了国内外著名高校、著名学者的精品课程, 课程内容涉及多个领域、多种学科甚至是多种语言。

我国在线课程的建设经历了跨越式的变革。2003年, 我国开启国家级精品课程项目, 2010年教育部获批了3910门国家精品课程; 2011年开启中国大学视频公开课、精品资源共享课, 截至2013年教育部公布了992门视频公开课, 2884门资源共享课。2017年, 国家对在线课程进行了审查和评价, 截至2020年, 认定3000余门国家精品在线开放课程, 其中不仅有经过2个教学周期检验的课程, 而且还将在新冠肺炎期间开展大规模在线教学中作出重要贡献的原2017年和2018年的课程也纳入了名单。在我国在线开放课程建设与应用取得

## \*基金项目

2019年度江苏高校哲学社会科学项目“成果导向理念下的在线课程学习效果评价研究”(编号: 2019SJA0116)。

2021年江苏省高等教育学会江苏高校“教学研究”项目南京林业大学课程思政建设探索与实践(编号: 2021JSY076)。

一定成绩的基础上,教育部首次正式推出了490门“国家精品在线开放课程”。2022年前启动“双万计划”,计划建设1万门国家级和1万门省级线上线下混合式一流课程,建立MOOC学分认定制度,推动优质资源在全国乃至全球开放共享。

在线课程的质量高低直接决定在线学习效果,学习者的高效果评价依托于高质量的在线课程。随着“互联网+”时代下大数据应用和数据分析的逐步成熟,在线课程的活动质量测评变得越来越技术化与工具化。

在国外,可汗学院基于信息跟踪技术研发的“数字化学习仪表盘”,可以对在线课程的学习过程进行精密追踪,记录大量的学生学习信息和情境,最终以数字和图标的方式呈现,成了教育大数据的可视化工具;普渡大学研发的“课程信号系统”可以依据学生过程性的活动信息对线上课程的学习结果进行判断并预警。

在国内,一方面,在遴选在线课程时仍然习惯于以专家为主导、基于标准的分析方法,但专家评审从本质上是专家通过以学生的身份来模拟上课,参与和体验线上课程的学习效果,某种程度上背离了学生的真实情况,存在一定的元认知“失真”。由于学习者在课程偏好、学习习惯、学习目的等方面存在多样性差异,因此通过这种方式遴选出来的课程存在“盲区”,无法精确反应学生的学习效果。另一方面,则是基于学习者的调查分析,例如“MOOC学院”从学习者的知识点、教师参与、趣味性、课程设计和难度等多重感知视角设计了课程质量分析方法。但这种方式得到的效果评价颗粒度比较大,多数以整个课程内容作为观测点,缺乏针对性。

## 二、马尔科夫模型的建立

### 1. 基本理论

设随机过程 $\{X_n, n \in \mathbb{N}\}$ , 如果对任意的 $n \in \mathbb{N}$ 及状态 $i_0, i_1, \dots, i_{n+1}$ , 有 $P\{X_{n+1}=i_{n+1}|X_0=i_0, X_1=i_1, X_2=i_2, \dots, X_n=i_n\}=P\{X_{n+1}=i_{n+1}|X_n=i_n\}$ , 则称 $\{X_n, n \in \mathbb{N}\}$ 为马尔科夫链。

若马尔科夫链的条件概率与当前时刻 $n$ 无关, 则称条件概率 $P^{(n)}=P\{X_{m+n}=j|X_m=i\}$  ( $i, j \in S, m \geq 0, n \geq 1$ ) 为马尔科夫链的 $n$ 步转移概率, 相应的称 $P^{(n)}=(p_{ij}^{(n)})=P^n$ 为 $n$ 步转移概率矩阵。

当 $n=1$ 时,  $p_{ij}^{(1)}=p_{ij}$ ,  $P^{(1)}=P$ .

规定 $p_{ij}^{(0)}=\begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$

$n$ 步转移概率 $p_{ij}^{(n)}$ 指的是系统从状态 $i$ 出发经过 $n$ 步转移

到状态 $j$ 的概率, 它对中间 $n-1$ 步转移经过的状态无要求。

### 2. 模型建立

在线课程的教学是一个较长的过程, 一个学期的授课可以看成若干个课堂时段, 每个课堂时段又可以分为若干个时间点<sup>[3]</sup>。学生在每个时间点的学习状态和获得的成就作为状态空间的子集, 每个状态都是具有随机性的, 假设教师授课方式、节奏把控等不发生较大改变的前提下, 前一个时间点的学生学习状态到下一个时间点的学生学习状态发生变化, 反映了教师这段过程中的教学效果, 学生的每次状态转移后的结果也只会和当前状态有关, 且学生在疫情之下也都是第一次全部经历线上教学, 排除了学生的基础差异性的影响。

本文以南京林业大学疫情期间开设的高等数学在线课程为样本, 大二年级普通班随机抽取不同老师授课的2个平行班级(1班和2班)共60人, 将其期中考试和期末考试的高数成绩作为统计数据, 利用马尔科夫链模型进行量化处理。期中考试作为初始状态, 期末考试作为转移后的状态。同时, 发布调查问卷, 针对线上课程的授课方式、影响因素等情况进行分析, 共回收有效问卷60份。最终学生期末考试的成绩进行“标准化”处理, 分为5个等级: 一等(100-90分)、二等(89-80分)、三等(79-70分)、四等(69-60分)、五等(59-0分)。

计算1班、2班期末考试平均成绩分别为76.85分和74.32分。单纯由均分来评判1班教学效果大于2班是不公正的。按照现在的状态, 继续教学后, 比较未来发展的趋势。计算1班、2班在期中考试和期末考试中由状态 $i$ 转移到状态 $j$ 频数, 得到其频数转移矩阵和对应的转移概率矩阵<sup>[4]</sup>。

由一步转移矩阵得到平稳分布, 将上述值 $P$ 分别带入平稳方程:

用Matlab解得极限向量为:  $\pi_1=(0.12, 0.24, 0.36, 0.2, 0.08)$ ,  $\pi_2=(0.25, 0.31, 0.29, 0.08, 0.06)$ 。分别选取95, 85, 75, 65, 55作为一等到五等的评价指标, 求得1班和2班的综合指标分别为:

$$M_1=0.12 \times 95+0.24 \times 85+0.36 \times 75+0.2 \times 65+0.08 \times 55=76.2$$

$$M_2=0.25 \times 95+0.31 \times 85+0.29 \times 75+0.08 \times 65+0.06 \times 55=80.35$$

由此可见, 综合水平 $M_1 < M_2$ , 在考虑原有基础差异情况下, 2班教学质量优于1班。

### 3. 问卷分析

为及时、系统地了解学生在线学习的情况以及教师线上授课的情况, 本次研究设计了师生版调查问卷。基于两个班的调查情况, 学生自身有一定的线上学习经历, 表示接触过线上课程的占比有75.51%。在学生的网络自主学习中, 观看

名校有关的课程视频排在首位,完成相应的课后作业排在末位。由此看出,线上自主学习如果没有硬性作业完成度的要求,很难形成完整的闭环链<sup>[5]</sup>。

经过疫情期间线上授课,学生们在填空式的问卷中互动活跃,提出不少建设性意见,一班更多学生表达了希望教师多互动的建议,例如“理科思辨要求较高的课程,老师直播授课语速快,学生记笔记跟不上”;“希望可以安排习题课讲解”;“希望老师增加与学生连麦互动的环节”;“下一堂课的重点难点提前告知学生,方便预习、复习”等实质性的问题。

#### 4. 独立性检验

由于1班和2班学生的成绩数据是样本数据,是整个学校总体学习高等数学的学生的代表,具有随机性,故需要用独立性检验的方法确认在考虑原有基础差异情况下,2班教学质量是否确实优于1班。根据问卷调查,1班教师互动多于2班,则以1班数据为例,关于线上教学中教师与学生之间有所互动,是否能提高学生对课堂内容的投入,可得如下教师与学生互动与学生投入之间的关系:教师互动多,学生投入状态好,关系值为35(a);教师互动多,学生投入状态不好,关系值为3(b),总计关系值38(a+b)。教师不互动多,学生投入状态好,关系值为8(c),教师不互动,学生投入状态不好,关系值为13(d),总计关系值21(c+d)。学生投入状态好时,教师状态总值44(a+c),学生投入状态不好时,教师状态总值16(b+d)。教师与学生互动与学生投入总计60(n=a+b+c+d)。

假设教师互动和学生投入课堂效率没有关系,相对独立为 $H_0$ ,两个分类变量相互有关为 $H_1$ , $|d-b|$ 越小,则教师互动和学生投入的关系越弱。

$$\text{根据公式 } K^2 = \frac{n(d-b)^2}{(a+b)(c+d)(a+c)(b+d)}$$

上述表格中数据代入公式得, $K^2 \approx 19.839$ ,在 $H_0$ 成立的情况下,根据独立性检验的临界值得: $P(K^2 \geq 10.828) < 0.001$ ,即这个概率非常小只有0.1%,因此可以有99.9%的把握认为1班的教学质量优于2班是因为教师互动的多,与前面求得的综合水平 $M_1 < M_2$ 相呼应。

#### 三、结论

基于马尔科夫链模型的教学效果评估方法,关注了评价依据的客观化、多元化,不再只局限于学生最终的成绩来评定一个老师上课的好坏<sup>[6]</sup>。同时,本文在测算模型的基础上对独立性进行了检验,将教师的“教”与学生的“学”相互

结合,使得教学评价不再是单一的结果性评价,而是更关注教师的教学过程性评价;不仅仅是主体评价,也考虑到了客体分析。

和传统课堂评价的主体分析不同的是,线上线下混合式教学的主体多了一个平台大数据。在“互联网+”的大数据时代,线上线下混合式教学是基于在线课程或SPOC,有20%-50%学生线上自主学习的教学模式。因此,混合式教学的课堂评价一部分依赖于平台的大数据分析,一部分来自学生和督导教师的课堂评价。平台大数据算法主要包括链接分析法、内容分析法、日志分析法。利用大数据算法对课程的教学绩效和学习绩效进行分析。

#### 四、措施

对比线上线下混合式教学的课程,传统授课方式的课程建设与在线课程建设在课程设计上也有共性也有区别。共同点在于师生需要有有效的互动,线上课程授课时设计的线上互动答疑、课堂小测验的答案统计排名等,如果熟练地运用教学用具在课堂中的使用频率,会让课堂效果翻倍。线下课程授课时学生的分组讨论、师生角色转换等,可以凝聚课堂人心。但在实施过程中,因为学习主体和学习客体的不同,具有差异性,同时还要考虑线上课程平台的实用性、技术支持以及技术人员的反馈时效等问题。老师的教学设计不仅仅要考虑课程内容,还需要考虑“互联网+”时代下线上线下教学方法的改革、学生获取知识能力和主动性的差异等等;信息化时代带给师生的不仅仅是信息获取的便捷和课程内容的丰富性,更重要的是要考虑在线课程的设计是否适合学生对知识点的掌握、线上与学生的教学互动是否及时、视频学习内容更新是否有阶段性等等;对学习效果评价的设计主要包括在线考试、随堂小测验、课后习题作业等指标。本课题探索建设基于Markov链的数据化评价,并对在线课程建设和教学效果进行系统性的评价,从而进行了成效分析。

#### 参考文献

- [1]张金桥.论高校教师的教学元认知能力及其培养[J].华南师范大学学报(社会科学版),2005(2):155.
- [2]范岩,马立平.利用马尔可夫链的高校教师教学质量模糊综合评价方法[J].数学的实践与认识,2017,47(4):77-82.
- [3]刘鲁文,陈兴荣,何涛.基于马尔科夫链的教学效果评估方法[J].统计与决策,2014,15(3):93-94.
- [4]董建国,惠淑荣,陈忠维.数学教育评价的马尔科夫链模型应用[J].沈阳农业大学学报(社会科学版),2010,12(4):468~471.