DOI:10.19789/j.1004-9398.2016.02.005

基于马尔科夫链的校本成绩评价合理性的探索*

景春丽 章建石
(教育部考试中心,北京 100084)

摘 要

本文利用马尔科夫链的预测模型来探索校本成绩评价的合理性,方法是将校本成绩和高考成绩均转换成 Z 分数并将 Z 分数划分成六个区间,统计各区间的人数,利用马尔科夫模型分别预测校本成绩及高考成绩在各区间的概率,并对其结果进行对比分析. 结果表明马尔科夫预测模型能比较准确的预测校本成绩的状态,也能够在一定程度上预测高考成绩的状态,但是预测能力非常有限,并探讨了预测误差比较大的原因.

关键词: 马尔科夫链,校本成绩评价,高考成绩,Z 分数.
中图分类号: O213.9, G449.7

1 校本成绩的概念及其使用价值

随着教育评价多元化的发展,建立健全和改进教学评价体系,已经成为迫在眉睫的重要任务[1]. 本文利用马尔科夫链的预测模型对校本成绩评价的合理性进行了探索.

高中的各阶段性测试、期中考试、期末考试等成绩称之为校本成绩. 在现行高校招生制度中,高中教师对学生的校本成绩的评价在录取中很难发挥实质性的作用,招生决策过度依赖某一次考试的结果,其效度是值得怀疑的.

校本成绩由于在评价维度、评价的时间跨度和频度、评价者信度等方面的优势,理应纳入高校选拔的考查范围,这方面欧美国家的高校积累了丰富的经验[2]. 不过,迄今为止,在国内各阶段的入学考试中,校本成绩被使用的情形很少. 如果能找到一种机制,对学生的校本成绩进行合理的校验,基于学生的平时成绩和老师对学生的评价来进行选拔,会有助于消解“一考定终身”、“应试教育”等违犯教育规律和人才选拔规律的做法.

2 模型与方法

2.1 马尔科夫链预测模型

马尔科夫链是马尔科夫(A. A Markov 俄国数学家)提出的一种随机时间序列,它在将来取什么值只与它现在的取值有关,而与它过去取什么值无关,即无后效性. 具备这个性质的离散随机过程,称为马尔科夫链.

2.2 样本描述

本文选取某重点高中高一、高二期中期末四次考试数学成绩作为校本成绩,高考成绩作为统一考试成绩. 样本采集涉及五次不同的考试,为了将五次考试转化成统一的标准并具有可比性,本文将五次成绩均转换成 Z 分数,并将 Z 分数从高到低分成 $(-3,-1]$,$(-1,-0.5]$,$(-0.5,0]$,$(0,0.5]$,$(0.5,1]$,$(1,3]$ 六个区间.

2.3 计算过程

利用状态转移概率和系统的初始状态可以建立马尔科夫预测模型[3],公式如下:

$$S(k+1) = S(k) \cdot P = S(k-1) \cdot P^2 \quad (k = 0,1,2,\cdots) \qquad (1)$$

本文中将高一期中、高一期末、高二期中、高二期末、高考作为不同的状态,从一个状态到另一个状态的变化具有稳定的趋势,这种转移依赖于学生目前的状态,与以往的状态无明显或直接的关系,具有无后性,所以教学过程可以看成一个马

收稿日期: 2014-01-11
* 基金项目: 全国教育考试专项立项课题成果(GFA111003)

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

尔科夫链.

表 1　高一、高二期中期末及高考成绩 Z 分数在各区间的人数

Z 分数区间	高一期中	高一期末	高二期中	高二期末	高考
(-3,-1]	27	22	30	27	25
(-1,-0.5]	18	19	20	19	7
(-0.5,0]	28	28	32	35	30
(0,0.5]	30	37	21	34	45
(0.5,1]	30	37	36	22	40
(1,3]	30	20	24	26	16
总计	163	163	163	163	163

3　结果与分析

3.1　利用高一期中、期末的校本成绩状态预测高二期中的校本成绩状态

高一期中、高一期末、高二期中三个状态的时间间隔是一个学期,将高一期中考试的成绩作为初始状态,记为第一个时刻,将高一期末考试的时间作为第二个时刻,计算高一期中到高一期末的状态转移概率,通过状态转移概率预测高二期中考试成绩在各区间的人数的比例. 初始概率为初始状态各区间学生人数与总人数之比.

初始概率:

$$S(0)=(0.17\quad 0.12\quad 0.17\quad 0.18\quad 0.18\quad 0.18)$$

表 2　高一期中到高一期末的成绩 Z 分数所在区间的转移情况

Z 分数区间	(-3,-1]	(-1,-0.5]	(-0.5,0]	(0,0.5]	(0.5,1]	(1,3]
(-3,-1]	10	4	3	3	5	2
(-1,-0.5]	4	3	6	3	2	0
(-0.5,0]	3	8	4	8	5	0
(0,0.5]	3	2	10	9	5	1
(0.5,1]	1	1	3	8	12	5
(1,3]	1	1	2	7	7	12

高一期中到高一期末时成绩 Z 分数的转移概率矩阵为:

$$P=\begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{pmatrix}$$

$$=\begin{pmatrix} 0.37 & 0.15 & 0.11 & 0.11 & 0.19 & 0.07 \\ 0.22 & 0.17 & 0.33 & 0.17 & 0.11 & 0 \\ 0.11 & 0.29 & 0.14 & 0.28 & 0.18 & 0 \\ 0.10 & 0.07 & 0.33 & 0.30 & 0.17 & 0.03 \\ 0.03 & 0.03 & 0.10 & 0.27 & 0.40 & 0.17 \\ 0.03 & 0.03 & 0.08 & 0.23 & 0.23 & 0.40 \end{pmatrix}$$

根据初始状态向量和状态转移概率预测高二期中考试成绩在各区间人数的比例,高一期中的状态经过两次转移达到高二期中的状态.

由公式(1)可预测高二期中的状态向量:

$$S(2)=S(1)\cdot P=S(0)\cdot P^2$$

$$=(0.17\quad 0.12\quad 0.17\quad 0.18\quad 0.18\quad 0.18)\cdot \begin{pmatrix} 0.37 & 0.15 & 0.11 & 0.11 & 0.19 & 0.07 \\ 0.22 & 0.17 & 0.33 & 0.17 & 0.11 & 0 \\ 0.11 & 0.29 & 0.14 & 0.28 & 0.18 & 0 \\ 0.10 & 0.07 & 0.33 & 0.30 & 0.17 & 0.03 \\ 0.03 & 0.03 & 0.10 & 0.27 & 0.40 & 0.17 \\ 0.03 & 0.03 & 0.08 & 0.23 & 0.23 & 0.40 \end{pmatrix}^2$$

$$=(0.13\quad 0.12\quad 0.19\quad 0.24\quad 0.22\quad 0.10)$$

表 3　高二期中考试成绩在各区间的人数比例的预测值与实际际值的对比分析

Z 分数区间	(-3,-1]	(-1,-0.5]	(-0.5,0]	(0,0.5]	(0.5,1]	(1,3]
实际值	0.18	0.12	0.2	0.13	0.22	0.15
预测值	0.13	0.12	0.19	0.24	0.22	0.1
绝对误差	0.05	0	0.01	0.11	0	0.05
相对误差	0.28	0.00	0.05	0.85	0.00	0.33
相对误差均值	1.51					

表 3 可以看出通过高一期中、高一期末预测高二期中的状态,预测值和实际值的误差整体比较小,相对误差值为 1.51,只有 Z 分数在(0,0.5] 区间的绝对误差和相对误差比较大之外,其他五个区间,误差都比较小,预测结果还比较准确.

3.2　利用高一、高二期末校本成绩状态预测高考成绩状态

高一期末、高二期末、高考三个状态的时间间隔是一个学年,将高一期末考试成绩 Z 分数在各区间人数的比例作为初始状态,初始概率为各区间学生人数与总人数之比,将高二期末考试的成绩 Z 分数在各区间人数的比例作为第二个状态,计算高一期末到高二期末的状态转移概率.

初始概率为:

$$S(0)=(0.13\quad 0.12\quad 0.17\quad 0.23\quad 0.23\quad 0.12)$$

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 4 高一期末到高二期末的成绩 Z 分数所在区间的转移情况

	(-3,-1]	(-1,-0.5]	(-0.5,0]	(0,0.5]	(0.5,1]	(1,3]
(-3,-1]	11	2	7	1	1	0
(-1,-0.5]	6	5	4	4	0	0
(-0.5,0]	3	5	6	5	6	3
(0,0.5]	2	4	8	9	6	8
(0.5,1]	4	1	9	12	5	6
(1,3]	1	2	1	3	4	9

高一期末到高二期末时成绩 Z 分数在不同区间的转移概率矩阵为:

$$P=\begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{pmatrix}$$

$$=\begin{pmatrix} 0.50 & 0.09 & 0.31 & 0.05 & 0.05 & 0 \\ 0.32 & 0.26 & 0.21 & 0.21 & 0 & 0 \\ 0.11 & 0.18 & 0.21 & 0.18 & 0.21 & 0.11 \\ 0.05 & 0.11 & 0.22 & 0.24 & 0.16 & 0.22 \\ 0.11 & 0.03 & 0.24 & 0.32 & 0.14 & 0.16 \\ 0.05 & 0.1 & 0.05 & 0.15 & 0.2 & 0.45 \end{pmatrix}$$

根据初始状态向量和状态转移概率预测高考的状态向量,高一期末的状态经过两次转移达到高考的状态.

由公式(1)可预测高考成绩的状态向量:

$$S(2)=S(1)\cdot P=S(0)\cdot P^2$$

$$=(0.13 \quad 0.12 \quad 0.17 \quad 0.23 \quad 0.23 \quad 0.12)$$

$$\begin{pmatrix} 0.50 & 0.09 & 0.31 & 0.05 & 0.05 & 0 \\ 0.32 & 0.26 & 0.21 & 0.21 & 0 & 0 \\ 0.11 & 0.18 & 0.21 & 0.18 & 0.21 & 0.11 \\ 0.05 & 0.11 & 0.22 & 0.24 & 0.16 & 0.22 \\ 0.11 & 0.03 & 0.24 & 0.32 & 0.14 & 0.16 \\ 0.05 & 0.1 & 0.05 & 0.15 & 0.2 & 0.45 \end{pmatrix}^2$$

$$=(0.17 \quad 0.13 \quad 0.21 \quad 0.19 \quad 0.14 \quad 0.16)$$

表 5 高考成绩在各区间的人数比例的预测值与实际值的对比分析

Z 分数区间	(-3,-1]	(-1,-0.5]	(-0.5,0]	(0,0.5]	(0.5,1]	(1,3]
实际值	0.15	0.04	0.18	0.28	0.25	0.1
预测值	0.17	0.13	0.21	0.19	0.14	0.16
绝对误差	0.02	0.09	0.03	0.09	0.11	0.06
相对误差	0.13	2.25	0.17	0.32	0.44	0.6
相对误差均值	3.91					

表 5 可以看出通过高一期末和高二期末的状态预测高考的状态,预测高考值和实际高考值的绝对误差在 0.02 ~ 0.11 之间,相对误差在 0.13 ~ 2.25 之间,相对误差均值为 3.91,两种误差都比较大,说明预测值与高考实际值差别比较大.

4 结论与讨论

根据分析结果,基于校本成绩能比较准确的预测下一个状态的校本成绩,校本成绩也能够在一定程度上预测学生统一考试的成绩,但是预测能力非常有限. 说明学生校本成绩与学生统一考试成绩之间还存在很大差距,这与更多其他因素有关. 首先,校本成绩考试的考查范围、命题难度与统一考试存在较大的差异. 其次,在我国高中教育中,课程教学大都在高二就结束,高三主要用于复习备考. 在这一学年中,由于学生态度的改变、学习技巧的获得、加上在学校学习强度的加大,学生的学习状态会发生很大的改观,学生的学科能力在这一年里一般都会发生较大的变化,从而使得校本成绩与统一考试成绩产生偏离,最终导致预测值与实际值偏差比较大.

参 考 文 献

[1] 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(3):93 - 94.

[2] 章建石. 后高考时代看美国多元化招生[N]. 中国教育报,2011 - 8 - 24005.

[3] 程伟. 马尔科夫模型在大学生转专业趋势预测中的应用[J]. 宿州学院学报,2014,29(3):112 - 114.

[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.

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Explore the Rationality of the School Evaluation Based on Markov Chain

Jing Chunli　Zhang Jianshi

(National Education Examinations Authority, Beijing　100084)

Abstract

The rationality of the school evaluation scores is explored by the forecast model of Markoff chain in this paper. The school evaluation scores and college entrance examination scores are transformed into z-score which are divided into six intervals. The number of each interval are counted. The probabilities in each interval of the school evaluation z-score and the college entrance examination z-score are predicted by Markoff model which are compared with the actual condition. The results show that the school evaluation state can be predicted accurately but the college entrance examination. The reasons of the big prediction error are discussed.

Key words: Markoff chain, school evaluation scores, college entrance examination scores, z-score.

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net