

工科大学数学成绩与过往数学学习情况的有序逻辑回归和因子分析

张 伟, 李秀森*

山东理工大学, 数学与统计学院, 山东 淄博
Email: 1270386459@qq.com, *leexiusen@163.com

收稿日期: 2021年7月15日; 录用日期: 2021年8月9日; 发布日期: 2021年8月16日

摘 要

通过对山东理工大学在校学生发放调查问卷, 收集学生过往成绩及平时学习情况的相关数据, 将收集信息数据化处理后, 把数据录入Spss进行因子分析。将收集数据中第二学期高数成绩划分为优秀、良好、不及格三个层次, 通过有序逻辑回归分析建立模型, 通过分析结果确定各变量对第二学期高数成绩的影响。

关键词

过往数学成绩, 平时学习情况, 因子分析, 有序逻辑回归分析

Ordered Logistic Regression and Factor Analysis of the Engineering College Students' Past Learning Performance on Their Mathematics Performance

Wei Zhang, Xiusen Li*

School of Mathematics and Statistics, Shandong University of Technology, Zibo Shandong
Email: 1270386459@qq.com, *leexiusen@163.com

Received: Jul. 15th, 2021; accepted: Aug. 9th, 2021; published: Aug. 16th, 2021

Abstract

Through the questionnaire distributed to the students in Shandong University of Technology, we

*通讯作者。

collected the relevant data of the students' past achievements and their daily learning. After the data was processed, the data were recorded in spss for factor analysis. The second semester high scores were classified into three levels: excellent, good and failing. The MODEL was established by logistic regression analysis, and the influence of each variable on the second semester high scores was determined.

Keywords

Past Math Scores, Usual Studies, Factor Analysis, Ordered Logistic Regression Analysis

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

当前国内对大学生成绩的关注度越来越高,方向也越来越广,裘晓岚、陈兰清、肖蓬[1]等对高校学生数学学习中男女生学习水平及男女生差异进行了分析研究,邱女[2]研究了性别、宿舍环境、文理性质等因素对大学生成绩的影响,杨金梅、李瑾、邱宗明[3]研究了智力因素、自制力、学习动机等方面对学生成绩影响,张志红、耿兰芳[4]针对学习表现、专业的喜好程度、考试态度,以及课堂交流和讨论学习等方面进行研究分析。本文主要研究过往数学成绩和平时学习情况对大学生未来数学成绩的影响,以山东理工大学能源动力、土木工程等专业的学生为调查对象,设计调查问卷,回收有效问卷 119 份。

2. 因子分析

将收集的问卷进行数据化处理后,将所得数据录入 spss,利用 spss 进行因子分析。如表 1 如示,近似卡方值为 952.819,显著性趋近于 0 小于 0.05,表中 KMO 的值为 $0.811 > 0.7$,说明变量间的偏相关性较强,所以适合做因子分析。

Table 1. Tests on KMO and Bartlett

表 1. KMO 和 Bartlett 的检验

	取样足够度的 Kaiser-Meyer-Olkin 度量	0.811
	近似卡方	952.819
Bartlett 的球形度检验	df	190
	Sig.	0.000

根据特征值大于 1 及碎石图分析可知,第一公因子对变量的解释度最高,第五个公因子后的因子特征值均小于 1,对变量的解释程度很低,第一个公因子的方差贡献率为 31.498%,前五个公因子累计方差贡献率为 60.395%,所以前五个公因子代表了所收集数据的大多数信息,故本次因子分析应该提取五个公因子合适,因子分析结果较理想。第一公因子主要包括能跟上老师提问思路、听讲专注程度、老师授课内容理解程度、上课使用电子设备、课后温习、高考数学成绩等指标情况,第二公因子主要包括检查作业情况、自主完成作业情况、作业完成度、报名模拟考试、保存作业情况,第三公因子包括第一学期高等数学成绩,是否有考研计划,恪守课堂规范情况,第四公因子包括复习笔记,错题集,第五公因子包括主动完成作业,做练习题,学习时间等情况。

根据成分得分矩阵列出因子得分方程

$$F_1 = 0.806 \text{ 跟上老师提问思路} + 0.786 \text{ 听讲专注程度} + 0.769 \text{ 老师授课内容理解程度} + 0.750 \text{ 上课使用电子设备} + 0.593 \text{ 课后温习} + 0.577 \text{ 主动预习情况} + 0.479 \text{ 高考数学成绩} + 0.441 \text{ 检查作业情况} + 0.546 \text{ 自主完成作业情况} + 0.542 \text{ 作业完成度} + 0.118 \text{ 报名模拟考试} - 0.008 \text{ 保存作业} + 0.072 \text{ 第一学期高数成绩} + 0.061 \text{ 是否有考研计划} + 0.230 \text{ 恪守课堂规范情况} + 0.076 \text{ 复习笔记} + 0.233 \text{ 错题集} + 0.170 \text{ 主动完成作业} + 0.301 \text{ 做练习题} - 0.345 \text{ 学习时间}$$

同理计算出 F2~F5, 根据因子得分结果分析, 学生的课堂学习情况是影响同学成绩最主要的因素, 其次作业及练习情况对学生成绩也有比较突出的影响效果, 所以大学生在制定学习计划时应注重把握课堂的学习时间, 认真听取老师的授课内容, 且过往成绩的优异程度对未来学期成绩影响不显著, 只有认真听课才是提高成绩的关键方法。

3. 有序逻辑回归分析

首先将收集的问卷中学生第二学期的高数成绩划为优秀, 良好, 不及格三个层级, 其中 80~100 为优秀, 记为 1, 60~80 为良好, 记为 2, 小于 60 为不及格, 记为 3, 其次对分类变量进行哑变量处理, 将跟上老师提问思路、听讲专注程度、老师授课内容理解程度、上课使用电子设备、课后温习、高考数学成绩、检查作业情况、自主完成作业情况、作业完成度、第一学期高等数学成绩, 恪守课堂规范情况, 可命名为课后温习、设为协变量, 是否考研, 自觉完成作业情况, 保存作业, 错题集, 复习笔记情况, 报名模拟考试, 做练习题等设为因子。

由表 2 可知容忍度均大于 0.1, 膨胀因子小于 10, 不存在多重共线性。

Table 2. Coefficient table

表 2. 系数表

模型	共线性统计量	
	容差	VIF
(常量)		
1		
高考数学成绩	0.681	1.468
主动预习情况	0.660	1.514
听讲专注程度	0.306	3.268
老师授课内容理解程度	0.319	3.135
恪守课堂规范情况	0.601	1.664
跟上老师提问思路	0.311	3.216
上课使用电子设备	0.505	1.980
课后温习	0.428	2.339
作业完成度	0.300	3.339
自主完成作业情况	0.255	3.927
检查作业情况	0.265	3.776
第一学期高数成绩	0.687	1.456
考研	0.689	1.452
自觉完成作业	0.675	1.482
保存作业	0.572	1.748
错题集	0.726	1.378

Continued

	笔记	0.792	1.262
1	参加模拟考	0.679	1.473
	做练习题	0.691	1.447

由表 3 可知: 平行线检验中的卡方为 15.537, $p = 1$, 大于 0.05, 接收原假设, 满足平行线检验, 可以使用有序逻辑回归分析。由模型拟合信息表知显著性为 0.001 小于 0.05, 表示本次研究的变量中至少有一个的系数不是 0, 所以模型成立。

Table 3. Parallel line check

表 3. 平行线检验

模型	平行线检验 ^a			
	-2 对数似然值	卡方	df	显著性
零假设	101.404			
广义	85.867 ^b	15.537 ^c	59	1.000

零假设规定位置参数(斜率系数)在各响应类别中都是相同的。

a) 联接函数: Logit。

b) 在达到最大步骤对分次数后, 无法进一步增加对数似然值。

c) 卡方统计量的计算基于广义模型最后一次迭代得到的对数似然值。检验的有效性是不确定的。

$$\ln\left(\frac{\gamma_1}{1-\gamma_1}\right) = -1.399 + 0.017\text{高考数学} + 0.22\text{第一学期高考数学成绩} \\ - (1.114\text{考研} - 2.880\text{保存作业} + \dots - 1.636\text{作业批改情况})$$

$$\ln\left(\frac{\gamma_2}{1-\gamma_2}\right) = 4.663 + 0.017\text{高考数学} + 0.22\text{第一学期高考数学成绩} \\ - (1.114\text{考研} - 2.880\text{保存作业} + \dots - 1.636\text{作业批改情况})$$

高考数学成绩和第一学期高等数学的成绩的回归系数分别为 0.017 和 0.22, 但并未呈现出显著性($p > 0.05$), 意味着高考数学成绩及第一学期高等数学成绩并不会对第二学期的高数成绩产生影响关系, 听讲专注程度理解程度, 课堂规范, 跟上老师提问思路, 上课使用电度, 课堂内子设备情况存在显著性($p < 0.05$), 上述因素对第二学期高数成绩的影响具有统计学差异。

根据参数估计值表格可列方程上述的两个方程只在阈值有差别。由此能够算出 γ_1 , γ_2 , 以此计算第二学期的高数成绩获得优秀, 良好, 不及格的概率, 进一步对第二学期的高数成绩进行分析和判断。将一号同学的各项数据代入, 计算得出 $\gamma_1 = 0.75$, $\gamma_2 = 0.24$, $\gamma_3 = 0.01$, 结果解释为该同学第二学期高数成绩为优秀的概率为 75%, 成绩为良好的概率为 24%, 不及格的概率为 1%, 综合分析后得出第一名学生第二学期的高数成绩等级为优秀。

部分预测结果表 4 中, EST1_1, EST2_1, EST3_1 分别代表第二学期高数成绩优秀, 良好, 不及格的概率, 其中第一行 EST1_1 概率最大, 为 0.75 对 PER_1 为 1 (优秀), 此模型准确的预测了第二学期的高数成绩水平。

统计 OR 值计算: 听讲专注程度理解程度, 课堂规范, 跟上老师提问思路, 上课使用电度, 课堂内子设备情况对应的 p 值小于 0.05, 上述因素对第二学期高数成绩的影响具有统计学差异, 其中听讲专注程度为 98 的较听讲专注程度为 30 同学优秀的倾向优秀的 122,352.517 倍, 课堂内容理解程度为 98 的较课堂内容理解程度为 50 倾向优秀的 33,810.220 倍, 课堂遵守规则情况为 98 的相比遵守规则为 90 的倾向

优秀 0.108 倍, 跟上老师提问思路为 98 的较跟上老师提问思路为 90 倾向优秀的 10.435 倍, 上课使用电子设备情况为 98 的较上课使用电子设备情况为 80 倾向优秀的 0.017 倍。根据表 5 求出模型的预测准确度为 $(48 + 46 + 3) \div 118 = 94.025\%$ 。

Table 4. Prediction result

表 4. 预测结果

EST1_1	EST2_1	EST3_1	PRE_1
0.75	0.24	0.01	1
0.65	0.34	0.01	1
0.27	0.67	0.07	2
0.67	0.31	0.01	1
0.67	0.32	0.01	1
0.06	0.66	0.28	2

Table 5. Table of predicted response categories

表 5. 预测响应类别表

第二学期高数成绩*预测响应类别交叉制表						
			预测响应类别			合计
			优秀	良好	不及格	
第二学期高数成绩	优秀	计数	48	10	0	58
		第二学期高数成绩中的 %	82.8%	17.2%	0.0%	100.0%
	良好	计数	8	46	0	54
		第二学期高数成绩中的 %	14.8%	85.2%	0.0%	100.0%
	不及格	计数	0	3	3	6
		第二学期高数成绩中的 %	0.0%	50.0%	50.0%	100.0%
合计	计数	56	59	3	118	
	第二学期高数成绩中的 %	47.5%	50.0%	2.5%	100.0%	

由上述的分析可知, 听讲专注程度理解程度, 课堂内容理解程度对第二学期的高数成绩的影响尤为突出, 其次上课能跟上老师的提问思路的对听课效果有较好影响, 遵守课堂规则和不使用电子设备的同学成绩也相对较好, 综上所述, 若学生想要提高自己的学习成绩就应该抓紧课堂的时间, 把课堂所学到的东西融会贯通, 紧跟着老师的授课思路和授课节奏, 同时要加强自己的自律自控能力, 不要在课堂时间开小差导致自己跟不上教学内容。

4. 结论

两种分析法的研究结果中, 听讲专注程度, 课堂内容理解, 课堂规范, 回答跟上思路, 上课使用电子设备情况均为影响第二学期高数成绩的重要因素, 综合分析结果表明, 学生如果想要提升自己第二学期的成绩, 就应该在课堂上做到认真听讲, 把课堂所学的内容融会贯通, 相比于课后的努力, 课堂对促进学生学习成绩的效果更大, 抓住课堂的时间加以利用就可以最好的提高自己的成绩, 而高考成绩及第一学期的高数成绩对第二学期的成绩无明显的影响, 所以同学们不要因为自己曾经的成绩优异而放松自己的学习态度, 尤其不应该因为自己过往的成绩优异而不听取老师的授课。

参考文献

- [1] 裘晓岚, 陈兰清, 肖蓬. 高师学生数学学习状况分析[J]. 福建师范大学学报: 自然科学版, 2009, 25(2): 119-124.
- [2] 邱女. 影响大学生成绩的因素分析[J]. 中国市场, 2011(26): 245-246.
- [3] 李瑾. 影响大学生学习成绩因素的分析[J]. 天津成人高等学校联合学报, 2005, 7(3): 55-59.
- [4] 张志红, 耿兰芳. 学习态度对大学生学习成绩影响实证分析[EB/OL]. 中国大学教学, 2009. <http://www.doc88.com/p-7448796164596.html>