

高校学生成绩影响因素的研究

顾永勋

(对外经济贸易大学 北京市 100029)

[摘要]高等院校在考察学生学习成果时,常见以考试成绩作为量化指标。何为影响大学生成绩的主要因素,已成为教育工作者们所热衷的研究课题。本文收集了某高校2016级理学院7个专业的学生成绩,并设计调查问卷获取相应学生的基本信息,涵盖学习态度、人口统计量 and 家庭文化背景三个方面的九个因子。采用R语言编写软件代码进行数据的处理,运用描述性统计分析、主成分分析和多元线性回归的因子分析等方法,阐明影响学生成绩的主要因素。

[关键词]学习成绩;多元线性回归模型;主成分分析;因子分析

1. 引言

近些年国内高等院校不断扩招,使更多的学子有机会获得大学教育。但是,高等学校学生的能力参差不齐也是客观存在的问题。因此,教育界人士越来越关注如何提高教育质量及成果。考试成绩作为考核学习成果的量化指标,被广泛运用于优秀大学生评定以及向用人单位优先推荐等等。面对庞大的成绩数据,大多数教育工作者往往只能停留于对学生成绩进行简单描述统计,未能开展深入的挖掘,致使很多数据资源无法被充分利用。本文将采用多元统计的方法和多种数据挖掘的技术对高校学生的成绩进行深层次的剖析,通过研究影响高校学生成绩的影响因素,为学生日常的学习行为提供更有利的指导和帮助,从而提高教育质量。

2. 数据准备

2.1 考试成绩数据的收集与预处理

本研究选取某高校2016级理学院的7个专业,包括材料物理、电子科学、光电学、统计学、微电子、信息科学和应用物理的学生在2019年的考试成绩进行分析。在整理的过程中,如果部分同学缺少的成绩过多,有可能是出现了学籍异动,则删除这些同学的成绩信息。经过数据统计整理后,理学院共计有287名学生的成绩。由于理学院不同专业的课程有所不同,因此部分课程无法代替学生的整体成绩,因此选择平均成绩的方式代替学生的综合成绩。数据整理完成后,对高校学生的平均成绩进行描述性统计分析,得出学生的学习成绩分布特征如图1所示。

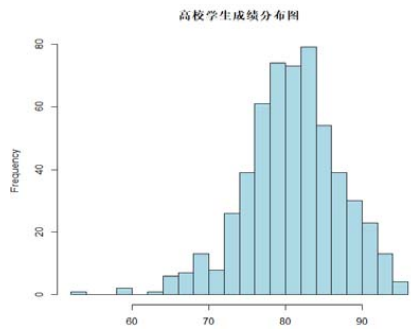


图1 平均成绩分布图

从图1可以看出,所有学生的平均成绩分布呈现近正态分布,平均值和中位值都是81分,最低分是53.8,最高分是95.3分,学生的总体成绩呈现良好。学生平均成绩按照90-100评定为优秀,80-89评定为良好,70-79评定为中等,小于70评定为差,经过这样划分,统计各个分数段内学生的频数,并制定表1。

表1 平均成绩评定频率表

		频数	百分比	累计百分比
有效	优秀	20	6.97%	6.97%
	良好	151	52.61%	59.58%
	中	100	34.84%	94.43%
	差	16	5.57%	100.00%
	合计	287	100%	

不同课程的难易程度和教师的教学方法存在差异,会对学生的成绩产生一定影响。针对不同专业学生的平均成绩根据t检验,电子科学专业的成绩与其他三个专业(材料物理,光电学和信息科学)的成绩存在显著性差异。因此后续研究仅考虑

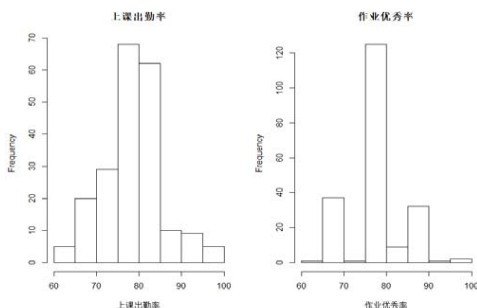


图2 上课出勤率与作业优秀率的分布图

除电子科学外其他六个专业学生(共计208人)的成绩做影响因素分析,这样就避免了专业不同对学生成绩存在影响。

2.2 影响因素数据的收集与预处理

经过文献调研,影响高校学生学习成绩的影响因素有很多,本文选取三大影响因素,分别是学习态度、人口统计学变量和家庭文化背景。每一影响因素中又涵盖了三个子影响因素。

学习态度的三个子影响因素包括每天的学习时间,上课出勤率,作业提交的优秀率。在学习的理论研究中,学习态度是影响高校学生学习的重要的非智力因素之一。图2展示了调查学生的上课出勤率与作业优秀率的分布,可以看出上课出勤率在75-85%之间分布最广,作业优秀率在75-80%之间分布最广。

人口统计学变量的三个子影响因素分别是性别,民族和居住地。通过调查问卷的方式对理学院的两百多名学生进行调查,除去电科专业的学生,剩余6个专业学生共计208人,调查学生的基本人口统计学变量如表2所示。

表2 人口统计学变量

人口特征变量		人数	百分比(%)
性别	男	127	61.06%
	女	81	38.94%
民族	汉族	178	85.58%
	少数民族	30	14.42%
居住地	城市	107	51.44%
	农村	101	48.56%

家庭文化背景信息影响因素分为是否父母离异,是否独生子女和家庭经济状况。是否父母离异考虑父母婚姻的变故对学生的心理产生影响,进而影响到成绩。是否是独生子女考虑家庭中小孩数量是否对学生的学习成绩产生影响。家庭经济状况考虑是否条件好的给学生更多的资源学习,或是家庭条件不好的学生是否更有动力学习。同样根据调查问卷获得208位学生的家庭文化背景信息如表3所示。

表3 调查学生的家庭文化背景变量

家庭文化背景		人数	百分比(%)
父母离异	是	40	19.23%
	否	168	80.77%
独生子女	是	176	84.62%
	否	32	15.38%
家庭经济条件	好	111	53.37%
	差	97	46.63%

3. 高校学生成绩的主成分分析

3.1 主成分分析各影响因素

本研究的九个影响学生学习成绩的影响因子(性别、民族、居住地、学习时间、上课出勤率、作业优秀率、是否父母离异、是否独生子女和家庭经济情况)经过主成分分析降维后,各个主成分因子可解释变量情况如图3碎石图所示,可以看出主成分1、2和3分别可以解释23.3%, 15.4%和13.9%的数据差异,依然不是很高,说明我们选取的9个影响因素不能很好的降维。

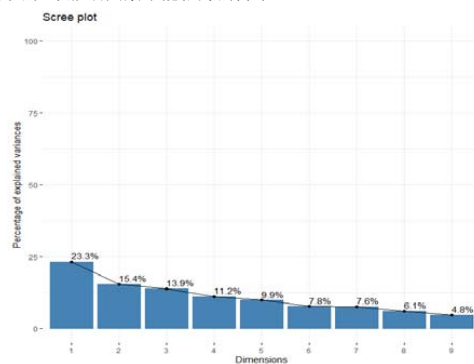


图3 主成分分析的碎石图

3.2 主成分分析结果描述

主成分分析各因子在主成分1和主成分2上的loadings数如表4所示，其中学习时间和上课出勤率对主成分1的贡献最大，作业优秀率对主成分2的贡献最大。可见在大的因子分类中学习态度对学习成绩起到关键作用。经过主成分分析降低维度后，数据采用主成分1和主成分2展示本研究的208位学生的成绩。如图4所示，主成分1和主成分2并不能很好的区分成绩评定为优秀、良好、中和差的学生。从另一方面也可以说明，评价学生的成绩不能由简单的几个变量来进行描述，需要又多变量进行分析。

表4 主成分分析各因子在主成分1和主成分2上的loadings数值

Table with 3 columns: Loadings, Factor1, Factor2. Rows include variables like 性别, 民族, 居住地, 学习时间, 上课出勤率, 作业优秀率, 父母离异, 独生子女, 家庭经济情况.

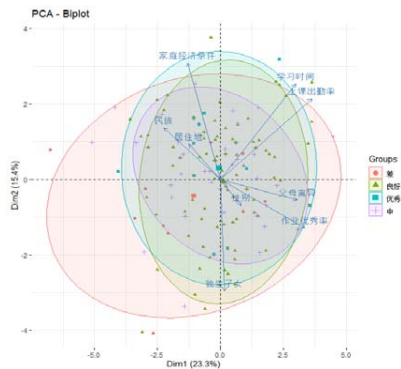


图4 主成分分析各影响因素

4. 高校学生成绩的因子模型分析

4.1 影响因素因子间的共线性检验

由于主成分分析后降维效果不好，因此因子分析中还继续使用9个影响因子进行分析。在建立多元回归模型前，需要检验各个因子之间是否存在共线性关系，也就是说是否存在相关性系数很高的两个因子。经过统计分析，本研究的九个影响学生成绩的因子（性别、民族、居住地、学习时间、上课出勤率、作业优秀率、

是否父母离异、是否独生子女和家庭经济情况）之间的相关性如图5所示，相关性最高的是学习与上课出勤率，相关系数为0.53，依然不高，因此本研究的9个因子之间不存在共线性关系，都可以用来分析建模。

4.2 多元回归模型的建立

本研究采用R语言对数据进行分析，首先分析9个因子的数据类型，需要将性别、民族、居住地、是否父母离异、是否独生子女和家庭经济情况转换为离散型因子变量(factor)，其他的如成绩、学习时间、上课出勤率和作业优秀率是连续性数字变量。然后建立多元线性回归模型，公式如下：

1m(成绩~性别+民族+居住地+学习时间+上课出勤率+作业优秀率+是否父母离异+是否独生子女+家庭经济情况)

4.3 多元回归模型的结果描述

统计分析结果如表5所示，显著的因子有6个(p<0.05)，其中最显著的两个因子是学习时间和作业优秀率(p<0.001)，其次是民族和上课出勤率(p<0.01)，最后是性别和父母是否离异(p<0.05)。居住地，是否独生子女和家庭经济条件对学生成绩的影响不显著。对于整个模型的拟合评价，p值显著，证明预测模型可信度高。修正后的可决定系数为0.4777，说明还存在其他影响考试成绩的因素。

结语

通过对学习态度，人口统计学变量和家庭文化背景这三大影响因素的主成分分析和因子分析结果说明，影响学生学习成绩最为重要的因素是学习态度，这种主观性因素的重要性要优于其他家庭和人口统计量等客观变量。而在九个子因素的影响程度分析中，学生每天的学习时间在多元回归模型中的，回归系数为3.96，且显著，是影响学习成绩所有因素中最为关键的因素。在主成分分析中，学生的上课出勤率对主成分1的贡献系数是0.916，排在第一位，对学生成绩的影响也是非常大。此外，作业优秀率，民族，性别和父母是否离异也对学生学习成绩的产生一定影响。

在分析了影响高校学生成绩的主要因素后，作者认为可以从学生、家长及学校三个方面出发寻求教育改进的措施。高等院校应加强师资队伍的建设，科学指定教学计划，让学生有充分的时间理解课程，有时间思考和实践，进而顺利且优质地完成学业。在家庭教育中，家长应该重视且尊重孩子的意愿，给予孩子必要的正确引导和帮助，并为学生营造交互式学习的家庭氛围。学业成绩取得进步的关键要素在于学生的自律性及能动性，正视自己的问题并能发挥自己的优势。坚守通过自身努力去改变自己命运的信念，强化自身的抗压能力，并培养自己对困难迎难而上的决心且善于向他人寻求必要的帮助。

参考文献

- [1]张莉.数据挖掘技术在高校学生成绩分析中应用的研究[D].中国石油大学,2009.
[2]金诗谱.基于属性相关性分析的高校学生成绩分析应用研究[J].西安文理学院学报:自然科学版,2018,v.21;No.84(01):35-38.
[3]曾志嵘,夏欧东,杜华,等.大学生学习成绩影响因素的分析和讨论[J].科教文汇,2006(8):31-34.
[4]张喜娟,丁钊鹏.因子分析在学生成绩影响因素调查结果中的应用[J].北京联合大学学报:自然科学版,2010(01):83-87.

作者简介:

顾永勤,男,对外经济贸易大学,统计学院在职人员高级课程研修班学员。

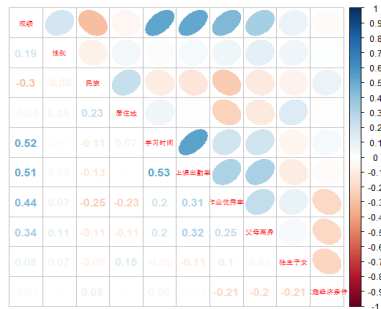


图5 各因子间的相关性

表5 多元线性回归模型的结果

Table with 6 columns: Variable, 回归系数, 标准误差, t-统计量, 显著性p, and significance markers. Rows include 截距, 性别, 民族, 居住地, 学习时间, 上课出勤率, 作业优秀率, 是否父母离异, 是否独生子女, 家庭经济条件, 拟合优度, 调整拟合优度, F-统计量, F检验显著性水平.