

# 基于因子分析和多元回归分析的大学生专业课成绩影响因素分析

曹凤彤, 李秀森\*, 王倩

山东理工大学数学与统计学院, 山东 淄博

Email: \*leexiusen@163.com

收稿日期: 2021年5月21日; 录用日期: 2021年6月9日; 发布日期: 2021年6月25日

## 摘要

通过大学生专业课成绩、相关基础课成绩、平时学习情况的相关分析得出: 在所有可能影响大学生专业课成绩的因素中, 是否有考研计划、课前预习情况、听课认真程度等七个因素与专业课成绩相关性较为显著的结论。为了减少存在多重共线性的变量并且不造成信息大量流失, 我们对变量进行了因子分析, 得到了因子分析模型与得分系数; 最后通过多元线性回归进行建模, 通过步进的方式筛选变量, 最终筛选出五个解释变量, 并得到线性回归方程: 大学生专业课成绩 = 33.824 + 0.263基础课期末考试成绩 + 0.131听课认真程度 + 2.060考前两周学习时间 + 0.076课前预习情况 + 0.068作业认真完成情况, 且这五个变量共同解释了因变量的59.9%。

## 关键词

专业课成绩, 因子分析, 多元线性回归分析

# Analysis of the Factors Influencing the Achievement of College Students' Professional Courses Based on Factor Analysis and Multiple Regression Analysis

Fengtong Cao, Xiuseen Li\*, Qian Wang

School of Mathematics and Statistics, Shandong University of Technology, Zibo Shandong

\*通讯作者。

文章引用: 曹凤彤, 李秀森, 王倩. 基于因子分析和多元回归分析的大学生专业课成绩影响因素分析[J]. 应用数学进展, 2021, 10(6): 2179-2185. DOI: 10.12677/aam.2021.106227

## Abstract

Through the correlation analysis of College Students' professional course performance, related basic course performance and usual learning situation, it is concluded that among all the factors that may affect college students' professional course performance, seven factors, such as whether there is a postgraduate entrance examination plan, the situation of preview before class, and the seriousness of attending class, are significantly correlated with the professional course performance. In order to reduce the multicollinearity of variables and not cause a lot of information loss, we make factor analysis on variables and get the factor analysis model and score coefficient; Finally, through multiple linear regression modeling and step-by-step selection of variables, five explanatory variables are finally selected, and the linear regression equation is obtained: College Students' professional course score = 33.824 + 0.263 final examination score of basic courses + 0.131 listening seriousness + 2.060 learning time two weeks before examination + 0.076 preview before class + 0.068 homework conscientiousness. These five variables together accounted for 59.9% of the dependent variable.

## Keywords

Results of Professional Courses, Factor Analysis, Multiple Linear 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

有关大学生专业课成绩及影响因素的调查研究开始于上世纪末。在国外有关本课题的研究中, Cuc Nguyen 和 Patrick Griffin 认为学生家庭财产指数、学生家长受教育年限等对学生学习成绩有显著影响[1], Hazura Mohamed 和 Noraidah Sahari 认为种族、学习项目是影响学生成绩的主要因素[2]。相较于国外, 国内在本课题方面的研究发展更为迅速, 所研究的影响因素种类也更加多样, 张志红等研究了学习态度与学习成绩的相关性[3], 路平等认为学学习动机通过学业自我概念对学习成绩有显著的正面影响[4]。本文立足于相关基础课成绩和平时学习情况两个方面研究它们与专业课成绩的相关性, 与以往研究相比, 本文主要考虑的是学业因素, 可以为想要提高专业课成绩的学生在基础课的学习以及平时学习习惯上给出可行性建议。

## 2. 对各影响因素的因子分析

对山东理工大学的学生进行问卷调查, 共收回调查问卷 203 份, 调查问卷的内容主要包括大学生专业课成绩、相关基础课成绩、平时学习情况三个部分。为了探究三者间的相关关系对问卷问题的选项进行赋值, 然后将数据导入 SPSS 软件中。为了找到影响大学生专业课成绩的因素, 先对调查问卷数据进行相关分析, 结果如下。

**Table 1. Correlation**  
**表 1. 相关性**

	1、文理科	2、您的基础课期末考试平均分是?	3、您的专业课平均分是?	4、是否有考研计划	5、考研科目里是否有数学	6、课前预习情况	7、听课认真程度
皮尔逊相关性	-0.038	0.671**	1	0.355**	0.280	0.584**	0.662**
3、您的专业课平均分是?	显著性(双尾)	0.590	0.000	0.000	0.647	0.000	0.000
	个案数	203	203	203	203	203	203
	8、课堂内容当时理解程度	9、课堂纪律遵守情况	10、回答老师提问,跟上老师思路	11、玩手机情况	12、复习课堂内容情况	13、作业认真完成情况	14、作业独立完成情况
皮尔逊相关性	0.698**	0.661	0.558**	0.589**	0.585**	0.646**	0.684*
3、您的专业课平均分是?	显著性(双尾)	0.000	0.012	0.000	0.000	0.000	0.004
	个案数	203	203	203	203	203	203
	15、是否能做到作业完成后自己认真检查一遍并改错	16、是否做到自觉做作业	17、对作业中的错题和重点题是否进行整理和保存	18、是否整理课程重点内容笔记和习题集	19、是否认真参加老师组织的模拟考试	20、考前是否做往届试题	21、考试前两周,用在本课程的学习时间是平时的几倍?
皮尔逊相关性	0.296	0.526*	0.246**	0.453*	0.113**	0.233**	0.456**
3、您的专业课平均分是?	显著性(双尾)	0.025	0.002	0.000	0.003	0.000	0.000
	个案数	203	203	203	203	203	203

\*\*在 0.01 级别(双尾), 相关性显著; \*在 0.05 级别(双尾), 相关性显著。

由表 1 所展示的各影响因素和专业课成绩的相关系数与显著性我们可以得出结论: 在所有可能影响大学生专业课成绩的因素中, 是否有考研计划、课前预习情况、听课认真程度、回答老师提问并跟上老师思路、复习课堂内容情况、作业认真完成情况以及考前在本课程的学习时间这几个因素与专业课成绩相关性显著。

为了消除变量之间的相关性以方便在下一步骤建立多元线性回归模型, 我们需要对数据进行因子分析来削减变量个数同时保证信息不会大量丢失。因子分析结果如下:

根据表 2, 由第一组数据项可知 21 个因子的方差贡献、解释原有变量总方差的百分比以及累计方差贡献分别为多少, 在初始解中共提取了 21 个因子, 原有变量的总方差均被解释, 累计方差贡献率为 100%; 其第二组数据项为提取了 4 个因子的情况, 4 个因子共同解释了原有变量总方差的 90.236%, 原有变量信息丢失较少, 因子分析效果较为理想。

**Table 2.** Explanation of total variance  
**表 2.** 总方差解释

成分	初始特征值			提取载荷平方和			旋转载荷平方和		
	总计	方差百分比	累积%	总计	方差百分比	累积%	总计	方差百分比	累积%
1	10.584	50.398	50.398	10.584	50.398	50.398	6.747	32.127	32.127
2	4.018	19.126	69.524	4.018	19.126	69.524	5.529	26.320	58.447
3	2.779	13.227	82.751	2.779	13.227	82.751	4.165	19.826	78.273
4	1.572	7.485	90.236	1.572	7.485	90.236	2.513	11.963	90.236
5	0.546	2.597	92.833						
6	0.290	1.385	94.218						
7	0.127	0.604	95.621						
8	0.112	0.533	96.154						
9	0.106	0.503	96.657						
10	0.097	0.462	97.119						
11	0.086	0.411	97.530						
12	0.078	0.369	97.899						
13	0.069	0.328	98.227						
14	0.062	0.295	98.522						
15	0.059	0.281	98.803						
16	0.051	0.245	99.048						
17	0.050	0.239	99.287						
18	0.046	0.219	99.506						
19	0.039	0.190	99.696						
20	0.036	0.173	99.869						
21	0.028	0.131	100.000						

提取方法：主成分分析法。

由表 3 的因子载荷矩阵，可以得到因子分析模型：

$$\text{作业独立完成情况} = 0.876f_1 + 0.043f_2 - 0.014f_3 - 0.075f_4$$

$$\text{作业认真完成情况} = 0.873f_1 - 0.016f_2 - 0.054f_3 + 0.053f_4$$

$$\text{课堂内容当时理解情况} = 0.865f_1 + 0.041f_2 - 0.037f_3 - 0.014f_4$$

.....

$$\text{考前两周学习时间} = 0.508f_1 - 0.020f_2 - 0.104f_3 - 0.634f_4$$

同时根据表 3 可以得到以下因子得分系数：

$$F_1 = 0.876\text{作业独立完成情况} + 0.873\text{作业认真完成情况} + 0.865\text{课堂内容当时理解情况} \\ + 0.846\text{听课认真程度} + \dots + 0.472\text{考前是否做往届试题} + 0.508\text{考前两周学习时间}$$

$$F_2 = 0.043 \text{作业独立完成情况} - 0.016 \text{作业认真完成情况} + 0.041 \text{课堂内容当时理解情况} \\ + 0.089 \text{听课认真程度} + \dots + 0.074 \text{考前是否做往届试题} - 0.020 \text{考前两周学习时间}$$

$$F_3 = -0.014 \text{作业独立完成情况} - 0.054 \text{作业认真完成情况} - 0.037 \text{课堂内容当时理解情况} \\ - 0.031 \text{听课认真程度} + \dots - 0.559 \text{考前是否做往届试题} - 0.104 \text{考前两周学习时间}$$

$$F_4 = -0.075 \text{作业独立完成情况} + 0.053 \text{作业认真完成情况} + 0.014 \text{课堂内容当时理解情况} \\ - 0.118 \text{听课认真程度} + \dots + 0.169 \text{考前是否做往届试题} + 0.634 \text{考前两周学习时间}$$

Table 3. Composition matrix

表 3. 成分矩阵

	成分			
	1	2	3	4
14、作业独立完成情况	0.876	0.043	-0.014	-0.075
13、作业认真完成情况	0.873	-0.016	-0.054	0.053
8、课堂内容当时理解程度	0.865	0.041	-0.037	0.014
7、听课认真程度	0.846	0.089	-0.031	-0.118
15、是否能做到作业完成后自己认真检查一遍并改错	0.820	-0.175	0.140	-0.076
9、课堂纪律遵守情况	0.819	0.107	-0.053	-0.042
11、玩手机情况	0.815	0.016	-0.055	-0.218
3、您的专业课平均分是?	0.796	-0.154	0.147	0.127
12、复习课堂内容情况	0.787	-0.209	0.058	-0.290
2、您的基础课期末考试平均分是?	0.779	-0.210	0.250	0.076
10、回答老师提问, 跟上老师思路	0.776	0.092	-0.065	-0.201
6、课前预习情况	0.735	-0.170	0.096	-0.370
16、是否做到自觉做作业	0.660	-0.131	0.007	0.352
18、是否整理课程重点内容笔记和习题集	0.613	-0.212	0.149	0.200
17、对作业中的错题和重点题是否进行整理和保存	0.606	-0.334	0.329	0.183
19、是否认真参加老师组织的模拟考试	0.562	0.231	-0.441	0.102
5、考研科目里是否有数学	0.497	0.739	0.193	-0.033
4、是否有考研计划	0.593	0.605	-0.101	-0.026
1、文理科	-0.087	0.568	0.610	0.186
20、考前是否做往届试题	0.472	0.074	-0.559	0.169
21、考试前两周, 用在本课程的学习时间是平时的几倍?	0.508	-0.020	-0.104	0.634

提取方法: 主成分分析法; a 提取了 4 个成分。

### 3. 各影响因素与专业课成绩的多元线性回归分析

通过相关分析与因子分析已经掌握了对专业课成绩影响较大的几个变量,为了分析各个变量间的统计关系,探究它们之间的数量变化规律,并且通过多元线性回归方程的形式对这种规律进行反应,现对调查问卷数据进行多元线性回归分析,结果如下。

由表 4 知,最终步进到模型中的变量有基础课期末考试平均分、听课认真程度、考试前两周用在该课程的学习时间、课前预习情况以及作业认真完成情况。(自变量步进的条件为自变量 F 的概率值  $\leq 0.05$ ,自变量被移除的标准为自变量 F 的概率值  $\geq 0.1$ 。)

**Table 4.** Variables entered/removed

**表 4.** 输入/除去的变量

模型	输入的变量	除去的变量	方法
1	2、您的基础课期末考试平均分是?	.	步进(条件:要输入的 F 的概率 $\leq 0.050$ ,要除去的 F 的概率 $\geq 0.100$ )。
2	7、听课认真程度	.	步进(条件:要输入的 F 的概率 $\leq 0.050$ ,要除去的 F 的概率 $\geq 0.100$ )。
3	21、考试前两周,用在本课程的学习时间是平时的几倍?	.	步进(条件:要输入的 F 的概率 $\leq 0.050$ ,要除去的 F 的概率 $\geq 0.100$ )。
4	6、课前预习情况	.	步进(条件:要输入的 F 的概率 $\leq 0.050$ ,要除去的 F 的概率 $\geq 0.100$ )。
5	13、作业认真完成情况	.	步进(条件:要输入的 F 的概率 $\leq 0.050$ ,要除去的 F 的概率 $\geq 0.100$ )。

a 因变量: 3、您的专业课平均分是?

由表 5 得多元线性回归方程:  $Y = 33.824 + 0.263X_1 + 0.131X_2 + 2.060X_3 + 0.076X_4 + 0.068X_5$  (Y 表示大学生专业课成绩,  $X_1$  为基础课期末考试平均分,  $X_2$  为听课认真程度,  $X_3$  为考前两周学习时间,  $X_4$  为课前预习情况,  $X_5$  为作业认真完成情况)。

**Table 5.** Coefficients

**表 5.** 系数

模型	未标准化系数		标准化系数	t	显著性	共线性统计	
	B	标准误差	Beta			容差	VIF
5、(常量)	33.824	3.920		8.629	0.000		
2、您的基础课期末考试平均分是?	0.263	0.068	0.260	3.870	0.000	0.452	2.211
7、听课认真程度	0.131	0.037	0.241	3.511	0.001	0.431	2.321
21、考试前两周,用在本课程的学习时间是平时的几倍?	2.060	0.707	0.153	2.915	0.004	0.738	1.355
6、课前预习情况	0.076	0.028	0.171	2.748	0.007	0.528	1.895
13、作业认真完成情况	0.068	0.032	0.148	2.104	0.037	0.413	2.421

a 因变量: 3、您的专业课平均分是?

## 4. 结论

由相关分析的结果可以得到结论:在所有可能影响到大学生专业课成绩的因素中,是否有考研计划、课前预习情况、听课认真程度、回答老师提问并跟上老师思路、复习课堂内容情况、作业认真完成情况以及考前两周学习时间这几个因素与学生专业课成绩相关性显著,且根据多元线性回归分析得到的回归方程:大学生专业课成绩 =  $33.824 + 0.263 \times$  基础课期末考试平均分 +  $0.131 \times$  听课认真程度 +  $2.060 \times$  考前两周学习时间 +  $0.076 \times$  课前预习情况 +  $0.068 \times$  作业认真完成情况。由此看来,一个高的专业课成绩离不开相关基础课的巩固学习、认真的听课态度、较长的学习时间、良好的课前预习习惯以及认真完成作业的习惯。

综上所述,要想得到一个良好的专业课成绩应该首先要重视基础课的学习,为将来学习专业课打好基础,其次还要课前认真预习、课上认真听讲、课后认真完成作业、考前两周争取在该学科多投入时间学习。另外,大家可以根据自己的时间、学习计划等在整理笔记、参加模拟考试、检查作业并改错、做往届试题等方面投入时间和精力。

## 参考文献

- [1] Nguyen, C. and Griffin, P. (2010) Factors Influencing Student Achievement in Vietnam. *Procedia-Social and Behavioral Sciences*, **2**, 1871-1877. <https://doi.org/10.1016/j.sbspro.2010.03.1001>
- [2] Mohamed, H., Ashaari, N.S., Judi, H.M., et al. (2012) Factors Affecting FTSM Students' Achievement in Statistics Course. *Procedia-Social and Behavioral Sciences*, **59**, 125-129. <https://doi.org/10.1016/j.sbspro.2012.09.255>
- [3] 张志红, 耿兰芳. 学习态度对大学生学习成绩影响实证分析[J]. 中国大学教学, 2009, 26(10): 102-104.
- [4] 路平, 刘聪颖, 夏福斌. 大学生学业自我概念、学习动机与学习成绩的关系[J]. 中国健康心理学杂志, 2016, 24(7): 1089-1092.