ISSN 1512-1801

ATTRIBUTION ANALYSIS OF PROBABILITY AND STATISTICS COURSE FOR ORIENTED NORMAL STUDENTS WITH LERNING DIFFICULTIES

[1]Ning Dong, [1*]Bo Yu, [2]Li Chen, [3]Wei Zhang

[1]School of Science, Hunan University of Technology
[2]School of Civil Engeneering[2] , Hunan University of Technology
[3]School of Foreign Language[2] , Hunan University of Technology
(* the corresponding author, boyu_hut@126.com)

Abstract:
As an important mechanism to supplement rural education resources, the oriented training of rural instructors is not only an appeal of the rural revitalization, but also an internal demand for balanced development of urban and rural education. In terms of the oriented students' academic performance of the course of probability and statistics, this paper first employs statistical methods to extract 12 principal components from 31 preset performance factors. Then the regression analysis is utilized to get the influence degree of the worked components on the sampled oriented students. The results show that self-management factors and have a positive impact on the academic performance of oriented students, while components of others' learning attitudes and their own professional and teachers' influence have a negative impact on their academic achievements.

Key-words: *Attribution; Oriented students; Factor analysis; Regression analysis;*

Introduction

Fairness and efficiency are the two major themes in the educational development. National long-term education in China requires regard "promoting fairness" as the basic policy of the educational development, and "improving quality" as the core task of education reform and development. However, under the background of the urban-rural development pattern, the imbalance of urban and rural education is obvious, which is manifested in the imbalance of urban and rural education investment, teaching resources and quality of instructors. Although with the steady progress of the urban-rural integrated development strategy, the imbalance of urban and rural education has been effectively alleviated, good rural instructors are still in urgent demand. Especially, such an imbalance is not manifested as a disparity in quantity, but the one in structure and quality (Ren & Chen 2020). Nowadays, the disparity in the quality of urban and rural instructors is mainly due to the fact that most of the teachers in rural schools are unqualified in their first academic qualifications, and their majors do not correspond to the courses taught. Therefore, the construction of the oriented rural instructors has become the most urgent task for the development of rural education. The implementation of the oriented rural instructor plan and the tailor-made targeted strategy of rural education are not only effective in alleviating the imbalance of urban and rural teachers, but also can help the balanced development of urban and rural education.

Under this background, Hunan University of Technology has joined the plan of balancing urban-rural development and has enrolled oriented rural students since 2018. After two years cultivation, these students have learned several fundamental courses, and correspondingly acquired certain teaching skills, taking on a trend of being integrated well into the undergraduate teaching plan. However, some oriented students, due to the relatively weak academic basis, are not able to catch up with the majority. This paper, in terms of one of their learned course --- Probability and Statistics, analyzes the possible reasons of the oriented students with learning difficulties.

ISSN 1512-1801

Probability and Statistics, based on the Advanced Calculus and Linear Algebra, has been a compulsory course for undergraduates. This course studies the quantitative relationship in random phenomena and is usually divided into two parts: the Probability Theory and the Mathematical Statistics. The former dwells on the regularity of random phenomena from the quantitative side while the latter utilizes the probability theory as the main mathematical tool to make inferences and predictions with collected random data. At present, it is widely penetrated and applied to various basic disciplines, engineering disciplines, economics, finance, management and other social disciplines.

Although the academic level of the textbook for oriented students is relatively interior to that for fully enrolled undergraduates. There are still many concepts and abstract theories seemly hard for them. From the perspective of instructors, the content is somewhat monotonous and, the teaching method, emphasizing symbolic calculation and calculation skills, usually ignores the practical ability. Meanwhile, the teaching goal focuses more on the training of logical thinking than stimulating the innovation of students. From the oriented students' perspective on the other hand, boring and abstract conceptual formulas make students "fear of difficulties" and lose interests. Besides, a large number of tedious and repetitive calculations also make oriented students fail to understand the ideas of probability statistics. All these result in the relatively high rate of failing of this course.

Attribution of academic performance refers to a cognitive activity in which students analyze their own learning behaviors based on test scores to infer the reasons for their academic performance (Xiao, et.al 2015). Seligman and Maier puts forward the concept of "learned helplessness", through the study of animal behavior (Seligman & Maier 1967). It means that an organism, when suffering successively from failures and frustrations and influenced by improper attribution and evaluation, feels losing control of everything and its psychological state becomes losing confidence in all behaviors (Zhang 1999). Given that the "learned helplessness" is unsolved, students with learning difficulties are produced. In view of the above reasons, performance attribution has become very important. At present, many quantitative and qualitative analysis of mathematical performance of college students have been made from different viewpoints. For example, multiple regression method was employed to find out relationship between mathematical achievements and self-concept, mathematics anxiety, recognition of mathematics value (Liu 2009); The causes and countermeasures of mathematics anxiety for non-mathematics major college students were investigated in (Peng et.al. 2011); The correlation between scores of college entrance examination and six personality factors were studied by (Wu et. al. 2011); Statistical methods were used to discuss the multi-relationships between the mathematics scores of freshmen and sophomores and individuals, families, society, and schools (Xiao et. al. 2015). However, there are few literatures on the attribution of the performance of oriented students with learning difficulties in Probability and Statistics. Therefore, this paper mainly focuses on the oriented students in "Probability and Statistics" to investigate and analyze the possible reasons for their poor performances.

Research Methodology

Research objects

A group test was used to conduct a cluster sampling survey for the oriented students of Hunan University of Technology who retook the course "Probability and Statistics" in the autumn of 2020. A total of 96 questionnaires were distributed and 102 valid questionnaires were returned. The questionnaire response rate was 90.2%. The subject of the survey involves the majors of the school elective courses such as Information and Computation, Economics, Civil Engineering, Biology & Chemistry, Applied Physics, and Communications & Computers, including sophomore, junior and

ISSN 1512-1801

senior students.

Study tool

By interviewing relevant teachers and students as well as referring questionnaire by (Xiao et. al. 2015), four levels of classification including individual, school, family, and society were determined. After discussing with instructors teaching "Probability and Statistics" in last semester, 6 and 3 secondary classification indexes were incorporated into "Individual" level and "School" level, finally resulting in the questionnaire covering 31 influencing factors that potentially affect the oriented students' academic performance. The following table lists all factors F_1-F_{31} and their corresponding descriptions.

Table 1
Potential influencing factors of oriented students' performance

Classification	Secondary Classification	Factor	Description
Individual	Course recognition	F_3	Helpless for individual major
	Basic level	F_1	Poor math level in the middle school
		F_2	Poor math level in the first year of university
	Psychological aspects	F_4	Fear of mathematics
		F_{28}	Nothing to fear
	Efforts	F_5	Frequent absences in class
		F_{11}	Frequently preview courses
		F_{12}	Frequently inquire of instructors
	Ordinary life	F_8	Abundant extracurricular learning activities
		F_9	Too many courses
		F_{25}	Addicted to internet games
	Self accomplishment	F_7	Bad memories of mathematical formula
		F_{10}	Not good at mathematics
School	Instructor's influence	F_{13}	Poor teaching
		F_{14}	Too difficult of exams
		F_{29}	More exercises in class
		F_{30}	Be more interesting
	Text book	F_6	Too abstract to understand
		F_{16}	Too boring
		F_{31}	Be more practical
	School spirit	F_{15}	Not strict in the final exam
		F_{17}	Classmates' learning attitude
		F_{19}	No strict supervision as in high school
		F_{22}	Greater impact of scholarship on students
		F_{23}	School disciplines
Family		F_{18}	Away from parental supervision
		F_{26}	Degree of family harmony
		F_{27}	Stress from parents
Society		F_{20}	Poor employment prospects
		F_{21}	Unpopular majors
		F_{24}	Impact of impetuous society

ISSN 1512-1801

Data processing

A class-based survey was conducted on the 31 factors influencing performance in the questionnaire. After the questionnaire was collected on the spot, 102 sample quantification tables were obtained but only 96 are valid. The software EXCEL was employed to count the sample data and SPSS 22.0 was utilized to analyze the derived statistical table. Concretely, the qualitative analysis of the questionnaire sample was first conducted, then the cluster analysis and factor analysis in the multivariate statistical analysis were applied to the obtained data. Finally, a regression analysis is performed on the relationship between the students' academic performances and the influencing factors.

Results Analysis and Discussion

Cluster analysis

(1) Analysis method

The average connection method (Zhu, 2012) which defines the square of the distance of different classes as the average of the square of the distance between two element, was employed to classify variables influencing performance in hierarchical cluster analysis.

(2) Analysis results

After performing average connection clustering on 31 factors, as well as synthesizing the practical significance, 3 categories were finally determined and the attribution of factors are listed as follows.

Table 2
Results of the cluster analysis for 31 factors

	Factors	Total Number
Category I	F_1, F_8, F_9, F_{11}, F_{15}, F_{26}, F_{27}	7
Category II	F_2, F_3, F_5, F_6, F_7, F_{14}, F_{16}, F_{17}, F_{18}, F_{19}, F_{25}, F_{28}, F_{30}, F_{31}	14
Category III	F_4, F_{10}, F_{12}, F_{13}, F_{20}, F_{21}, F_{22}, F_{23}, F_{24}, F_{29}	10

Factor analysis

The reducing-dimensionality technology combining cluster analysis and principal component analysis (Li 2013) is not only conducive to analyzing the meaning of principal components, but also can eliminate the multi-collinearity among variables (Mei et. al. 2006). Thanks to the KMO value of each category data exceeds 0.75, the factor analysis is remarkably proper in our analysis. Then the software SPSS 22.0 was exploited to the clustered three categories after extracting the principal components. According to the statistical principle that the eigenvalue of the correlation coefficient matrix is greater than 1, various factors were classified in the following table.

Table 3
Results of the analysis of principal components for 31 factors

	Principal components	Cumulative variance contribution rate
Category I	X_1, X_2, X_3	62.42
Category II	X_4, X_5, X_6, X_7, X_8, X_9	65.75
Category III	X_{10}, X_{11}, X_{12}	69.39

ISSN 1512-1801

Several linear relationships between the principal component and the potential factor are presented as below:

$X_1 = 0.032F_1 - 0.482F_8 + 0.548F_9 + 0.31F_{11} + 0.105F_{15} + 0.746F_{26} + 0.942F_{27}$;
$X_2 = 0.204F_1 + 0.374F_8 + 0.701F_9 - 0.421F_{11} + 0.581F_{15} + 0.682F_{26} - 0.8321F_{27}$;
…
$X_{12} = 0.432F_2 - 0.041F_4 + 0.041F_{12} - 0.041F_{14} + 0.041F_{21} + 0.041F_{28} + 0.413F_{31}$;

According to the principle that the greater the absolute value of the factor loading coefficient, the stronger the correlation between the factor and the corresponding factor, the relevant factors of each principal component are extracted, and the corresponding factors are named according to the nature of the extracted factors (see Table 3).

Table 4
Principal components and their name & description

principal component	Elements of Factors	Name of principal component
X_1	F_8, F_{26}, F_{27}	Daily life
X_2	F_9, F_{11}	Initiatively learning
X_3	F_1, F_{15}	Middle school mathematics foundation and examination room discipline
X_4	F_{18}, F_{19}, F_{25}	Self-management and other supervision
X_5	F_6, F_8, F_{16}, F_{28}	Difficult, boring and self-confidence
X_6	F_2, F_5, F_7	College mathematics foundation and effort factor
X_7	F_{29},F_{30}, F_{31}	Classroom application is less, uninteresting factors
X_8	F_{17}	Learning attitude
X_9	F_3, F_{14}	Not helpful for professional study and difficult examination
X_{10}	F_{20}, F_{21}, F_{22}, F_{23}, F_{24}	School and social impact factors
X_{11}	F_{10},F_{13}	Own profession and the instructor’s influence
X_{12}	F_4,F_{12}	Overcoming difficulties

ISSN 1512-1801

It is seen from the above table that the 12 extracted principal components have several notable features. The first is that the 12 components are extensively to cover all 31 pre-determined performance impact factors. The secondary is that the strong clusterity, 6 out of 12 principal components are connected with factors of the school level, 5 at the personal level, 1 at the family level, and 0 at the social level. The main impact of performance is attributed to the school level. In addition, the main influencing factors of each component have no overlap.

Together with the Table 1, one can further find that three factors of the social level are included in the same principal component, and the impact is relatively concentrated. It also shows that the factors at the social level are relatively independent, not being the main influence that affects the students' achievements with learning difficulties; Besides, the four factors at the teacher level in the secondary classification are scattered in four different principal components, showing a wide impact on students' academic performances. This indicates that the daily instructing activities are very important, and it is a key factor affecting the students' performances on probability and statistics.

Regression analysis of academic performance and principal components

In order to judge the strength of the 12 principal components on the results of the "Probability and Statistics" of oriented students, we used the final examination results of 112 students in the second semester retake classes of Hunan University of Technology from 2019 to 2020 to conduct a regression analysis between the students' academic achievements and the principal components. The regression model is set as a linear model:

$$Y = b_0 + b_1X_1 + \ldots + b_{12}X_{12} + \varepsilon$$

Among them, Y represents the scores of the sampled students, X_i (i = 1, 2,..., 12) is calculated by (1) to obtain the i-th impact component data of the sampled students. Again, SPSS 22.0 was employed to conduct the linear regression. With the backward regression method and deleting the independent variables, the regression equation is obtained as follows:

$$Y = 0.\ 201X_4 + 0.\ 303X_8 - 0.\ 421X_{11}$$

Obviously, this equation shows that three components of daily life, the learning attitude of others, and the influence of one's own profession and teachers have a greater impact on the performance of students in the retraining class.

Analysis of components in regression model

The principal component X_4 means that "self-management and other supervision", including extracurricular learning activities, self harmony, and ordinary self-supervision under proper pressure, impose deep affects on students' that students' daily life and learning activities. Another principal component X_8 is the learning attitude, including self-learning attitude and the impact of others' learning attitude on oneself. In our investigation, the learning attitude of classmates or roommates is found having a great influence on one's learning activities. Especially, students with poor self-control ability are more easily affected by bad habits. Finally, X_{11} is an influencing principal component of their own majors and teachers. Many students who retook the exam attribute their failure in the exams to instructors' poor teaching, which has a side effect on their learning.

If the influencing factor in Factor analysis is substituted into the derived regression model and sorted in order of the absolute value from the large to the small, the main influencing factors of the sampled scores could be extracted. There are 7 out of 31 principal components are greater than 0.35 in terms of absolute value: F_{17}, F_{26}, F_{27}, F_{29}, F_8, F_{13}, F_9. From the perspective of the composition of 7 main influencing levels, the social level accounts for 0, indicating that the sampled students are

ISSN 1512-1801

not concerned about the recognition of their majors and their employment prospects. In other words, factors from social level have a relatively weak impact on their performance. An interesting found is that there are respective two factors works for the individual level and the family level and these two levels had similar effects on the performance of the sampled students. The factors from school level accounted for four, including the learning attitude of others and the influence of the teacher's teaching. This indicates that the ability of self-control and independent learning of the sampled students is required to be strengthened, and the teaching behavior of instructors should also be changed reasonably.

Results & Discussion

According to the above analysis, the main factor affecting the performance of the oriented students with difficulty in probability and statistics lies in the school level, followed by the personal level. Especially, the learning attitudes of others and instructors' teaching behavior both have a relatively great impact on the sampled students' academic performances, reflecting their lack the ability of self-control and autonomous learning. As a contrast, the family level and social level give rise to little effects on their study achievements. Therefore, several measures might be put forward to increase oriented students' study ability.

Improvement of instructor's sensitivity and student's cognition

The course of probability and statistics has its own unique thinking feature, differentiating not only from the deterministic mathematical thinking method used in other disciplines, but also from and the traditional thinking methods which are too detailed and complicated to form a unified and coherent teaching cognition. The difference of students is not an obstacle to instruct, but a kind of valuable teaching resource. "Respect for the differences of students" should become the basic principle of teaching design (Zhong 2010). This requires instructors to make a series of various diagnoses, such as how to stimulate oriented students' cognitive conflicts in the course of teaching, how to discover students' cognitive impairments and help them in time, These are all indispensable thinking and practice to improve teachers' sensitivity.

Attention to students' life experience an d emotional experience

Oriented students with learning difficulties constitute a small part of the class group. The occurred cases of retaking this course make them doubting themselves. One of the most direct reasons is that they are lack of interest in learning, which might be caused by the dullness of teaching textbooks and the dull explanation of teachers. Therefore, instructors are supposed to stimulate students' interest in learning by altering the classroom to an environment full of life and emotion experiences. Life is the power source of education, the classroom is not only a knowledge classroom, but also an emotional classroom. Instructors should pay attention to introducing practical real-life cases of probability and statistics into the class and lifting the life meaning in the classroom. Meanwhile, the obscure mathematical language ought to be translated into passionate teaching words from real life and easy to be exploited by students. In this way can instructors stimulate students' motivation and desire to learn, and endow the classroom with emotional meaning.

Acknowledgments

This research is supported in part by 2020 Key Foundation of Hunan Education Department, 2020 General Foundation of Hunan Education Department.

ISSN 1512-1801

References

1. Ren S. & Chen Q. (2020). *The improvement of the public expenditure oriented training mechanism for rural teachers and risk prevention and control*. Jianghan Academic., 39 (3), 62-68.
2. Xiao J., Zhang L. & Shi E. (2015). *Statistical analysis of the influencing factors of mathematics achievement of college students of different grades*, Journal of Mathematics Education, 2015, 24(4), 53-56.
3. Seligman M.E & Maier S. F. (1967). *Failure to escape traumatic shock*. J. Exp. Psychol. 74: 1-9.
4. Zhang D. (1999). *Educational psychology*. People's education press.
5. Liu Z. (2009). *Research on the relationship between mathematics self-efficacy, self-concept and mathematics academic achievement of college students*. Journal of Mathematics Education, 18(6), 37-41.
6. Peng N. Liao S. &Chen Y. (2011) *Analysis of the causes of non mathematics majors' mathematics anxiety and countermeasures*. Journal of Mathematics Education, 20(3), 47-50.
7. Wu X., Hou X. & Zhu S. (2007). *The personality factors affecting the students' mathematics achievement*. Journal of Mathematics Education, 16(2), 68-70.
8. Zhu J. (2012). *Applied multivariate statistical analysis (2nd Edition).* Science Press, 2012.
9. Li H.(2013). *Application of principal component analysis method in analyzing direction college students' academic record*, Journal of Gansu Lianhe University, 27(6), 9-13.
10. Mei C.& Fan J. (2006) *Methods for data analysis*. Higher Education Press.
11. Zhong Q. (2010) *Research on classroom interaction: implications and topics*. Educational Research, 12(10): 73-80.

Article received 2020-10-26