Assessment of the Importance of Undergraduate Courses By Means of Multiple ANOVA

Mojieming Wei[1,a], ZiJun Mao[1,b], Mingxin Yang[1,c], Zhenjiang Wu[2,d], Yulong Gan[1,e], Liqun An[3,f,*]

kbei@mail.tust.edu.cn[a], zijunmao@mail.tust.edu.cn[b], 1719454398@qq.com[c], 2455584109@qq.com[d], 3427567870@qq.com[e], * Corresponding author:2395924462@qq.com[f]

College of Artificial Intelligence, Tianjin University Science and Technology, Tianjin, China[1]
College of Economics and Management, Tianjin University Science and Technology, Tianjin, China[2]
College of Sciences, Tianjin University Science and Technology, Tianjin, China[3]

Abstract. Basic courses at university are fundamental for establishing a robust basis in a college scholar's academic journey. In this context, the study of advanced mathematics has a crucial standing in several areas of inquiry. This study investigates the influence of elements including educators' proficiency and the administration's strategies on mathematics education's efficacy at the university level. The study investigates the impact of mathematics knowledge and skills on other university courses. The researchers adopted a multi-factor ANOVA statistical model to analyse the data. The findings indicate that the level of teacher instruction and the students' mathematical foundation have more significant effects than the level of college management. Therefore, it is recommended that the university improves the level of teacher and college management to facilitate students' holistic learning experience.

Keywords: chi-square test; multivariate ANOVA; polynomial fit

1 Introduction

College students frequently question the proportion of higher mathematics, linear algebra, and probability statistics covered in their courses. It is crucial to note that a positive mindset towards learning is essential for effective learning. In addition, what influence do mandatory public foundation courses have on both public foundation courses and professional courses? Facing unfamiliar and familiar subjects across different regions, disciplines, classes, teachers, and courses can significantly impact individual learning. Therefore, exploring influential factors of university foundational courses can aid students in boosting their academic performance while simultaneously raising the school's teaching standards. This paper will primarily utilise university mathematics as an example to analyse.

2 Methods

2.1 Factors affecting college mathematics

In order to explore the influence of college management level and teachers' teaching level on college mathematics learning, we find that the factors affecting college mathematics

EIMT 2024, March 29-31, Wuhan, People's Republic of China

DOI 10.4108/eai.29-3-2024.2347627

achievement mainly include college management level, teachers' teaching level and students' mathematics foundation. Therefore, the multi-factor variance model can be used to evaluate its influence [1]. The management level of the college can be reflected according to the mean score of college mathematics of the college students and the degree of improvement compared with the mathematics foundation. The teaching level of the teacher can also be reflected according to the mean score of college mathematics of all the students under the teacher and the degree of improvement compared with the mathematics foundation. On the premise of using Chi-square test to ensure that the data conform to the normal distribution [2], the variance analysis is carried out to better reflect the teaching level ability. Finally, multi-factor analysis is carried out to get the three influencing factors respectively on higher mathematics [3], linear algebra, probability and statistics, so as to get more appropriate results. The mathematical model is as follows:

$$y_{ijk} = \mu+\alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk} \tag{1}$$

$$\varepsilon_{ijk} \to N(0, \sigma^2) \tag{2}$$

$$\sum_{i=1}^{a} n_i\alpha_i = 0; \sum_{j=1}^{b} n_j\beta_j = 0; \sum_{k=1}^{c} n_k\gamma_k = 0 \tag{3}$$

If the factors considered are A and B, and there are a and b at each level, then the model of this problem is:

$$\begin{cases} y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \\ i = 1,2, \dots, a; j = 1,2, \dots, b; k = 1,2, \dots, m; \\ e_{ijk} i.i.d \sim N(0, \sigma^2) \\ \sum_{i}^{a} \alpha_i = 0, \sum_{j}^{b} \beta_j = 0 \\ \sum_{i}^{a} (\alpha\beta)_{ij} = 0, i = 1, \dots, a \\ \sum_{j}^{b} (\alpha\beta)_{ij} = 0, j = 1, \dots, b \end{cases} \tag{4}$$

The sum of squares of the total deviation $SS_T = \sum_i^a \sum_j^b (y_{ij} - \bar{y})^2$ is decomposed into:

$$SS_T = SS_A SS_B SS_{AB} SS_C \tag{5}$$

Among them, SST represents the total variation of the observation variable, SSA and SSB represent the variation caused by the independent action of control variables A and B, SSAB represents the variation caused by the interaction of control variables A and B, and SSe is the variation caused by random factors; the respective calculation formulae are as follows:

$$SS_A = \sum_{i}^{a} \sum_{j}^{b} \left(y_{ij} \bar{y}_i\right)^2 \tag{6}$$

$$SS_B = \sum_{i}^{a} \sum_{j}^{b} \left(y_{ij} \bar{y}_i\right)^2 \tag{7}$$

$$SS_C = \sum_{i}^{a} \sum_{j}^{b} \sum_{K}^{n_{ij}} \left(y_{ijk} - \bar{y}_{ijk}^{AB}\right)^2 \tag{8}$$

The value of Pr(F) is then obtained according to the hypothesis, which is as follows

$$\begin{cases} H_1: \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0 \\ H_2: \beta_1 = \beta_2 = \cdots = \beta_b = 0 \\ H_3: (\alpha\beta)_{ij} = 0, i = 1, \ldots, a; j = 1, \ldots, b \end{cases} \tag{9}$$

If different assumptions are valid, the corresponding F-value is:

$$F = \frac{MS_A}{MS_C} \sim F\left(a - 1, ab(n-1)\right) \tag{10}$$

$$F = \frac{MS_B}{MS_C} \sim F\left(b - 1, ab(n-1)\right) \tag{11}$$

$$F = \frac{MS_{AB}}{MS_C} \sim F\left((a-1)(b-1), ab(n-1)\right) \tag{12}$$

That is, if the value of Pr(F) is less than the significance level α, it means that the influence of this factor exists, otherwise it does not exist, and α=0.05 and α=0.01 are generally used to judge.

By analyzing the scores of each course in the data set, the following relationship chart can be obtained: the average scores of probability statistics and linear algebra are shown in Figure 1 and 2; the comparison between the scores of mathematics research and the mean values of university mathematics is shown in Figure 3; the comparison between probability statistics, linear algebra and mean values of each school is shown in Figure 4.

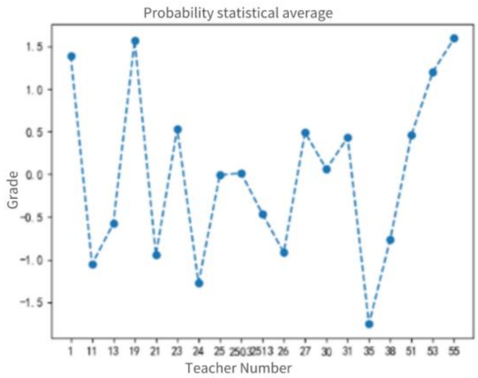

Figure 1. Probability statistics average grade comparison graph.

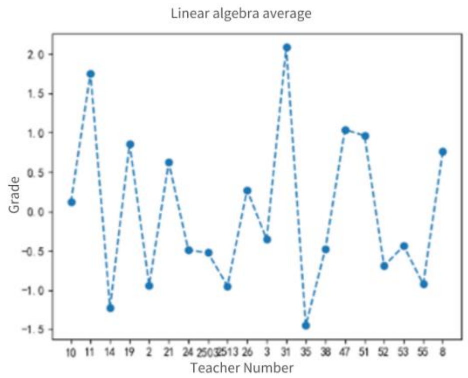

Figure 2. Linear algebra average comparison graph.

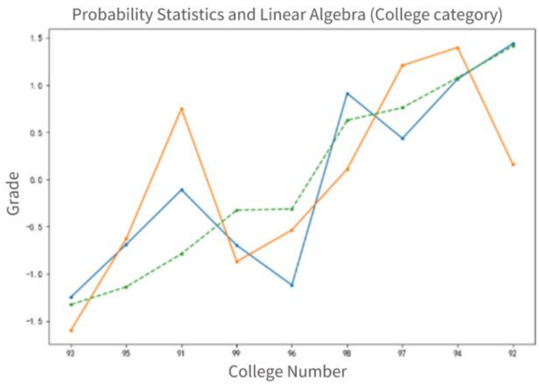

Figure 3. Graph of probability statistics and linear algebra (faculty category).

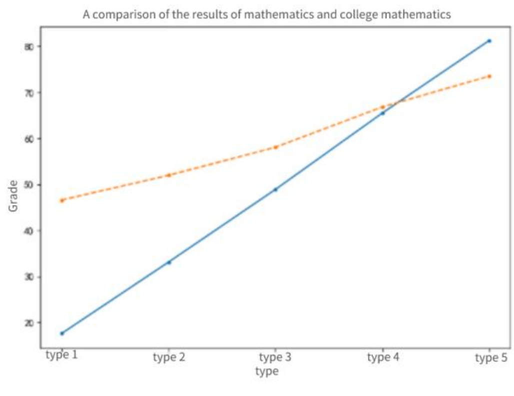

Figure 4. Comparison chart of math score and college math mean.

When assessing college mathematics, it's important to take into account both college mathematics and mathematics foundation scores. Thus, it's necessary to compare not only the average scores of both, but also the level of middle school mathematics improvement.

The ranking of relative improvement is:

15,58,7,45,18,2,28,8,40,13,22,6,9,46,36,57,34,1,25,4,44,42,30,2508,27,29,16,39,23,20,48

In terms of absolute ranking:

8,25,2,58,22,13,7,1,18,42,16,4,28,40,45,9,23,46,2503,6,57,39,48,36,20,27,4,15,34,30,29

It's evident from the above teacher ranking that students with higher teaching performance averages demonstrate greater improvement. In terms of factors affecting performance, teachers are a key player. Our preliminary analysis suggests that the following factors have the greatest impact in ranked order: teachers' teaching level, students' maths level, and college management level. This serves as a preliminary conclusion and provides a reference for subsequent variance analysis, eliminating major errors.

The influence of teachers is reflected in Table 1, which shows the variance and standard deviation for each subject.

Table 1. Variance and standard deviation of each subject.

Segment	Variance	Standard Deviation
Advanced mathematics	17.19770336025299	4.14701137691386
Upward appreciation	43.86409837154693	6.622997687720186
Probability statistics	16.95801193333019	4.118010676689679
Linear algebra	35.458629756330204	5.954714918140935

Under the influence of the college, the variance and standard deviation of each subject are shown in Table 2.

Table 2. Variance and standard deviation of each subject.

Segment	Variance	Standard Deviation
Advanced mathematics	12.015477294401133	3.46633485029954
Upward appreciation	11.825709815697941	3.4388529796573075
Probability statistics	18.490610247849443	4.300070958466783
Linear algebra	25.222725006927643	5.02222311401312
Mean values of probability statistics and linear algebra	8.593439811564899	2.9314569434949744

Based on the variance distribution for each factor, we can tentatively conclude that the overall fluctuation is greater when teachers are the influential factor in comparison to the college management. This is because variance indicates the degree of data fluctuation, leading us to understand that the teaching level of teachers has a higher impact on the data than college management. To achieve our goal, we measure disparities among educators in two categories: (1) the contrasts in absolute accomplishments; and (2) progress relative to secondary school. As we attach greater importance to the improvement aspect, the weightage is 0.35 and 0.65, the absolute achievement is designated as Si, and the relative enhancement is designated as Ci. Therefore, the formula for gauging the variance level encompasses:

$$U_{\text{score}} = 0.35(S) + 0.65(C) \tag{13}$$

Here is a multiple factor analysis of variance for the three subjects of advanced mathematics, probability statistics and linear algebra.

The influence of each factor on the higher mathematics subject at the university is finally obtained and shown in Table 3.

Table 3. The influence of each factor on higher mathematics in university.

	Degree of freedom	Sum of squares	Mean square sum	F value	PR (>F)
C(Senior math teacher last semester)	30	55446.48	1848.22	6.47	0.00**
C(Senior math teacher next term)	29	16011.74	552.13	1.93	1.984221e-03
C(college)	8	5988.94	748.62	2.62	7.361058e-03
C(Mathematical knowledge results)	42	14246.62	3392.06	11.87	0.00**
Residual	3489.0	996656.62	285.65	NaN	NaN

Based on the achieved probability p value, it is evident that all factors have contributed to learning higher mathematics in universities, as they all fall below the benchmark $\alpha= 0.05$. However, the teacher of higher mathematics and the mathematics foundation hold relatively high significance, passing the significance test of $\alpha=0.01$. This indicates a relatively considerable impact on higher mathematics in universities [4]. However, for the majority of students, the instructor for advanced mathematics is identical to the instructor for ordinary mathematics courses. Therefore, it can be inferred that the lecturer has a significant impact on advanced mathematics education in tertiary institutions.

Table 4 illustrates the effects of different factors of linear algebra on linear algebra subject matter.

Table 4. The effect of each factor on linear algebra.

	Degree of freedom	Sum of squares	Mean Square sum	F value	PR (>F)
C(Linear algebra teacher)	19.0	1.136037e+05	5979.144620	16.564146	0.00**
C(college)	8.0	1.906498e+04	2383.122101	6.602012	0.00**
C(Mathematical knowledge results)	42.0	8.092306e+04	1926.739539	5.337686	0.00**
Residual	3500.0	1.263392e+06	360.969080	NaN	NaN

In the provided data, all factors have a considerable impact on linear algebra and pass the $\alpha= 0.01$ significance test. However, the college's contribution is comparatively lower than the impact on advanced mathematics and still smaller than other factors.

Table 5 illustrates the influence of each probability statistics factor on linear algebra subjects.

Table 5. The influence of each factor on probability statistics.

	Degree of freedom	Sum of squares	Mean Square sum	F value	PR (>F)
C(Probability and statistics teacher)	18.0	63691.527810	3538.418212	25.283498	0.00**
C(college)	8.0	10492.341649	1311.542706	9.371529	6.9955 19e-13
C(Mathematical knowledge results)	42.0	31024.675154	738.682742	5.278201	0.00**
Residual	3490.0	488424.482642	139.949708	NaN	NaN

Based on the data presented above, the results are comparable to those found in linear algebra. Specifically, the influence of college on achievement is greater than that of advanced mathematics, but it is not as significant as the influence of teachers and one's own mathematical foundation. Finally, a ranking of the three factors that affect student achievement has been obtained: teacher teaching level > student mathematics foundation > college management level. However, for different mathematics subjects at the university level, the significance level of each factor varies. Nevertheless, the influence of teachers' teaching level and students' mathematics foundation is always high, whereas the influence of college management level becomes increasingly important with the progression of the semester and the variation of subjects.

3 Conclusions

After analysing the data set, it is evident that there is only a moderate correlation between college mathematics scores in general. It is also revealed that there are other factors affecting it. Through a detailed examination, including teachers' teaching level, college level, and students' mathematics level, it is evident that teachers' teaching quality has a significant impact on students' mathematics scores, while college level plays a less significant role. Therefore, enhancing the teaching proficiency of educators and providing students with comprehensive education will ultimately lead to an improvement in academic performance. During data processing, the overall distribution performance may get compromised, and hence, utilizing PCA and other data reduction techniques can assist in better summarizing the data and identifying significant influencing factors [5]. Additional evaluation approaches could be utilized to assign weights to the internal influencing factors of the data, including fuzzy evaluation [6].

References

[1] Fama E F. Multifactor portfolio efficiency and multifactor asset pricing[J]. Journal of financial and quantitative analysis, 1996, 31(4): 441-465.

[2] Tallarida R J, Murray R B, Tallarida R J, et al. Chi-square test[J]. Manual of pharmacologic calculations: with computer programs, 1987: 140-142.

[3] Dorčák P, Markovič P, Pollák F. Multifactor analysis of online reputation of selected car brands[J]. Procedia engineering, 2017, 192: 719-724.

[4] Farris J S, Kallersjo M, Kluge A G, et al. Constructing a significance test for incongruence[J]. Systematic biology, 1995, 44(4): 570-572.
[5] Roweis S. EM algorithms for PCA and SPCA[J]. Advances in neural information processing systems, 1997, 10.
[6] Icaga Y. Fuzzy evaluation of water quality classification[J]. Ecological Indicators, 2007, 7(3): 710-718.