



Research on graduation thesis data based on fuzzy comprehensive evaluation

Wei Tao, Jieyao Ling

Faculty of Mathematical and Applied Sciences, Nanjing Normal University Taizhou College, China
Email: 2798660689@qq.com

Abstract Graduation thesis (design) is an important index to measure the quality, level and efficiency of college education. In the current new situation of comprehensively improving the quality of undergraduate teaching, the competent department of education has issued documents, proposing to strengthen the management of the whole process of graduation project (thesis), strictly monitor the form, content and difficulty, and establish a random check mechanism. Improve the quality of graduation project (thesis) requirements. Each university revised and improved the regulations of graduation thesis management according to the requirements, and conducted research on the path to improve the quality of graduation thesis according to their different school conditions and learning conditions. This paper aims to study the data of graduation thesis system used by students and instructors. Firstly, the thesis process analysis system of PDCA management model is constructed by entropy value method, and then the fuzzy comprehensive evaluation matrix is constructed. Finally, according to the AHP method, the fuzzy comprehensive evaluation score is calculated to rank, and the influence degree of each link of the paper management on the paper assessment results is analyzed.

Keywords Graduation thesis data analysis; PDCA management system; AHP; Fuzzy comprehensive evaluation.

1. Introduction

In December 2020, the Ministry of Education issued a notice on the "Undergraduate Graduation Thesis (Design) sampling Inspection Measures (Trial)", which further strengthened the teaching supervision, evaluation and monitoring, and improved the basic level of undergraduate talent training. With the rapid development of science and technology, the wide application of ChatGpt will have a great impact on the graduation thesis review. Students over-reliance on natural language processing tools is bound to lead to a decline in the quality of graduate papers, which directly affects the quality of talent training. Under the social situation of fierce employment competition, the design and research of graduation thesis system is of great significance to promote the grid management of colleges and universities. Through the introduction of model driven and data analysis research methods to improve the optimization of graduation thesis system, it is convenient for universities to statistics and analysis of the implementation of graduation thesis in different stages, so as to further improve and perfect education supervision, evaluation and monitoring, in order to guide students to complete graduation thesis more actively and efficiently, and at the same time form a scientific basis for universities to formulate graduation thesis management system.



2. Research on data analysis of graduation thesis based on fuzzy comprehensive evaluation

2.1. Indicators of PDCA management system

PDCA management system is the Plan (Plan), implementation (Do), Check (Check), processing (Action) for short, is the total quality management work must go through the 4 stages.

This paper selects four stages, namely the planning stage (P stage), the implementation stage (D stage), the inspection stage (C stage) and the treatment stage (A stage). At the same time, in the established evaluation index system, each level index contains several secondary indexes. Including: the planning stage (P stage) of the advisor title factor, a secondary index; In the implementation stage (Stage D), there are 7 secondary indicators: the number of instructors, the number of submission of thesis topics, the number and proportion of submission of task papers, the number and proportion of submission of proposal reports, the number and proportion of submission of mid-term inspection, the number and proportion of subject modification and the overall evaluation score; A secondary index of the total number and proportion of finalized tests in the inspection phase (Phase C); The completion rate of the final draft of the paper in the processing stage (Stage A) is a secondary indicator. A total of 10 secondary indicators.

2.2. Indicator weights of the PDCA management system

The entropy weight method can be used to determine the weights of various indicators in the data analysis system. The entropy weight method reflects the importance degree of the same type of index by comparing the different relationships among each sample value in the index system of the same level. Entropy is a systematic and scientific measure. It is assumed that the smaller the entropy of an index of the same type is, the more information the index of the same type will provide, so it will have a better effect in fuzzy comprehensive evaluation and the greater the corresponding weight ratio.

First, calculate the entropy of the index:

$$Q_{ij} = \frac{M_{ij}}{\sum_{i=1}^m M_{ij}} \quad (1)$$

$$e = \frac{-1}{\ln(m) \sum_{i=1}^m Q_{ij} \ln Q_{ij}} \quad (2)$$

Where M_{ij} is the quantity value of a certain index, e_j is the entropy value of P_i

Then calculate the weight W_i of index M_j

$$W_j = \frac{1 - e_j}{\sum_{i=1}^n (1 - e_j)} \quad (3)$$

By combining the known data, the weights of each indicator are obtained as shown in Table 1:

Table 1: Thesis data analysis system A based on PDCA management model

Target layer	Criterion layer	First order weight	Index level	Second-order weight
Thesis data analysis system based on PDCA management model A	ProjectA ₁	0.12	Advisor title factorsA ₁₁	1
			Number of tutorsA ₂₁	0.102
			Topic submission timesA ₂₂	0.049
	ActA ₂	0.63	The number and proportion of assignment documents submittedA ₂₃	0.126
			The number and proportion of proposal reports submittedA ₂₄	0.138
			The number and proportion of mid-term inspection submissionsA ₂₅	0.266
			Number of subject modificationsA ₂₆	0.172
	CheckA ₃	0.16	Overall ratingA ₂₇	0.147
			The total number and proportion of finalized testsA ₃₁	1
	DealA ₄	0.09	Final draft completion rateA ₄₁	1



3. Comprehensive evaluation of graduation thesis data based on PDCA management system

Step 1: Determine the weight set of PDCA management system indicators

This paper first constructs a comprehensive evaluation matrix, assuming that the management system contains m row vectors, each row vector has n related evaluation indicators, then the initial comprehensive evaluation matrix X is constructed as follows:

$$X = [x_{i,j}]_{m \times n} \quad (4)$$

The specific processing process is as follows:

$$p_{i,j} = \frac{x_{i,j}}{\sum_{i=1}^m x_{i,j}} \quad (5)$$

The weight sets of PDCA management system indicators are as follows:

$$W = [w_1, w_2, \dots, w_n] \quad (6)$$

Step 2: Divide PDCA management model index level interval

With the help of the data analysis system of PDCA management model, the index value is divided into five different levels: worst, poor, medium, good and excellent, and according to the index, it is divided into four first-level indicators: planning stage, implementation stage, inspection stage and processing stage. The four phases are the planning phase (P phase), the implementation phase (D phase), the inspection phase (C phase) and the treatment phase (A phase).

Among them, in the graduation paper data analysis system, each first-level index contains several second-level indicators. In the planning stage (P stage), the title factor of the instructor is a secondary index; In the implementation stage (Stage D), there are 7 secondary indicators: the number of instructors, the number of submission of thesis topics, the number and proportion of submission of task papers, the number and proportion of submission of proposal reports, the number and proportion of submission of mid-term inspection, the number and proportion of subject modification and the overall evaluation score; A secondary index of the total number and proportion of finalized tests in the inspection phase (Phase C); The completion rate of the final draft of the paper in the processing stage (Stage A) is a secondary indicator. A total of 10 secondary indicators.

Step 3: Construct fuzzy evaluation matrix

This paper adopts the fuzzy comprehensive evaluation method to evaluate the graduation thesis data, regards the evaluation target as a data set composed of multiple fuzzy elements, then assumes the evaluation level of these fuzzy elements to form a fuzzy set, and then calculates the influence degree of a single fuzzy element on the evaluation level. Finally, according to the weight ratio of each fuzzy element in the evaluation target, Plug it into the formula and calculate the solution. By dealing with fuzzy evaluation objects with precise numerical calculation method, we can make scientific, reasonable and practical quantitative evaluation on the data containing fuzzy information.

Let U be data set $U = \{u_1, u_2, u_3\}$

$$u_1 = \{X_{11}, X_{12}, X_{13}\}, u_2 = \{X_{21}, X_{22}, X_{23}\}, u_3 = \{X_{31}, X_{32}, X_{33}\}$$

V is evaluation set $V = \{v_1, v_2, v_3, v_4\}$: v_1 means very good; v_2 means better; v_3 stands for general; v_4 means worse.

Construct evaluation matrix R_i :

$$R_1 = \begin{bmatrix} n_{111} & n_{112} & n_{113} & n_{114} \\ n_{121} & n_{122} & n_{123} & n_{124} \\ n_{131} & n_{132} & n_{133} & n_{134} \end{bmatrix}$$

$$R_2 = \begin{bmatrix} n_{211} & n_{212} & n_{213} & n_{214} \\ n_{221} & n_{222} & n_{223} & n_{224} \\ n_{231} & n_{232} & n_{233} & n_{234} \end{bmatrix}$$



$$R_3 = \begin{bmatrix} n_{311} & n_{312} & n_{313} & n_{314} \\ n_{321} & n_{322} & n_{323} & n_{324} \\ n_{331} & n_{332} & n_{333} & n_{334} \end{bmatrix}$$

Step 4: Calculate the fuzzy comprehensive evaluation vector according to AHP method

AHP method is adopted to obtain the weights of X layer to U layer as follows: (Weight of two level to one level index)

$$N_1 = (n_{11}, n_{12}, n_{13}) \quad N_2 = (n_{21}, n_{22}, n_{23}) \quad N_3 = (n_{31}, n_{32}, n_{33})$$

Let the first-order evaluation vector of U_i be B_i , so $B_i = N_i R_i = \{b_{i1}, b_{i2}, b_{i3}, b_{i4}\}$. In the secondary evaluation, consider each as $U_i (i = 1, 2, 3)$, U is a factor set, and the single factor evaluation matrix of U is

$$R = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix} \quad (7)$$

The weight of U layer to A layer is obtained by AHP method : $A = (a_1, a_2, a_3)$, and the second-order evaluation vector can be obtained

$$B = (b_1, b_2, b_3, b_4)$$

Normalize $B = (b_1, b_2, b_3, b_4)$, Then, the data results are added to take the average value. If the average value is greater than 0.5, the index obtained has a strong influence on the graduation thesis of students. On the contrary, the influence is weak.

Step 5: Improve the maximum membership principle

This paper assumes that the vector set of fuzzy comprehensive evaluation is $B = (b_1, b_2, b_3, \dots, b_n)$.

If $b_r = \max\{b_i\}$, The overall level of comprehensive evaluation belongs to grade r. And let's say that the vector we're building is $B = (0.563, 0.362, 0.070, 0.003, 0.002)$. Since the largest element in the vector is 0.563, the evaluation result is "good" according to the principle of maximum membership.

Meanwhile, the effectiveness indicators determined in this paper are defined as follows:

$$a = \frac{n\beta - 1}{2\gamma(n - 1)} \quad (8)$$

n is the number of unknowns in the normalized vector $B = (b_1, b_2, b_3, \dots, b_n)$; β is the maximum membership degree in $B = (b_1, b_2, b_3, \dots, b_n)$; γ is the second largest membership degree in $B = (b_1, b_2, b_3, \dots, b_n)$.

4. Data test and empirical analysis

According to the test criteria, the autocorrelation coefficient of order 0 is 1, and the autocorrelation coefficient of order k delay is 0 under certain conditions. In fact, the sample has limitations, the delay K-order autocorrelation coefficients are not all 0, as long as the value is near 0, it is considered that there is no autocorrelation. Because of the uncertainty, the autocorrelation coefficient is not exactly equal to 0, and in 95% of cases, the correlation coefficient is in a fixed range. If a sequence has more values of the autocorrelation coefficient outside the boundary, then the sequence is probably not a sequence.

After the initial data is processed by MATLAB, the data of the graduation thesis is tested by SPSS software. The obtained autocorrelation and partial autocorrelation images are shown in Figure 1 below:



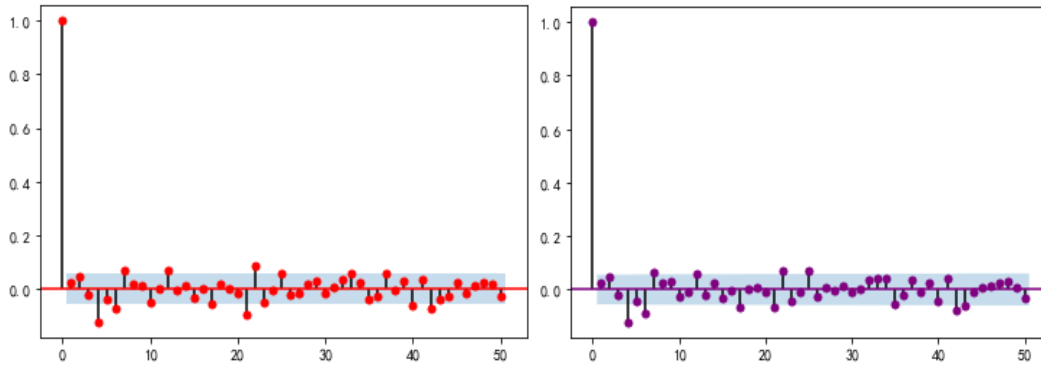


Figure 1: Autocorrelation and partial autocorrelation

All coefficients in Figure 1 are within the upper and lower critical values, that is, there is no significant difference from 0, so the selected fuzzy comprehensive evaluation model is more appropriate.

Through the simulation of actual data, The fuzzy comprehensive evaluation ranking of graduation thesis data index (1 to 10) is successively as follows: the number and proportion of submission of mid-term examination, the number of subject modification, the total evaluation score, the number of instructors, the title factors of instructors, the number and proportion of submission of proposal report, the number and proportion of submission of task book, the number and proportion of submission of thesis topic selection, the total number and proportion of final examination, and the completion rate of final thesis draft. The results are shown in Table 2:

Table 2: Comprehensive evaluation of graduation thesis data index

Indicators of PDCA management model	Score	Ranking
Final draft completion rate A_{41}	0.231	10
The total number and proportion of finalized tests A_{31}	0.311	9
The number and proportion of proposal reports submitted A_{24}	0.531	6
Number of tutors A_{21}	0.589	4
Number of subject modifications A_{26}	0.678	2
The number and proportion of mid-term inspection submissions A_{25}	0.741	1
Overall rating A_{27}	0.592	3
Advisor title factors A_{11}	0.511	5
The number of paper topic submission times A_{22}	0.426	8
The number and proportion of assignment documents submitted A_{23}	0.467	7

5. Conclusion and Suggestion

Conclusion

According to the results of comprehensive evaluation of graduation thesis data, the frequency and proportion of submission of mid-term examination are the most important to the completion of graduates' thesis. Secondly, the frequency of subject modification will also have a greater impact on the completion of graduates' thesis. The influence coefficients of the submission times and proportion of the proposal report, the number of instructors and the total evaluation score are all about 0.5, which is in the third echelon of affecting the graduation thesis. The influence coefficients of the title factors of instructors, the submission times of the thesis topic selection and the submission times of the task book and the proportion are about 0.4, which is in the fourth echelon. Finally, the completion rate of the final draft of the thesis is the least important for the completion of the graduate thesis.

Present situation and suggestions of graduation thesis

At present, the quality of undergraduate graduation thesis is affected by the problems such as insufficient time for students to complete the thesis, limited energy of instructors and outdated and absent management system. In addition, the lack of a more systematic scientific research training system to a certain extent on the quality of graduation papers also play a certain restriction. In the data analysis system of the graduation thesis system, the graduation thesis is affected by many aspects. Only by refining the existing graduation thesis management system and optimizing the reform of the graduation thesis management system can the quality of the graduation thesis be improved.



Judging from the evaluation results, the completion rate of the final draft of the thesis has little impact on the quality of the thesis, but the submission of the mid-term examination determines the quality of the graduation thesis to a large extent. Many graduates lack the concept of time for the completion of the thesis, while the mid-term examination of the thesis can timely and effectively help graduates grasp the degree of completion. In addition, for graduates who encounter problems in the research process and need to modify the research direction, The mid-term inspection can help them check the logic and science of the thesis, so as to avoid the phenomenon of poor quality of the thesis or the problem of the research direction in the graduation defense, and reduce the number of students who delay graduation and fail to graduate. In addition, the relevant factors such as the tutor and the number of thesis system tests have limited influence on the quality of the thesis. First of all, from the perspective of students, students are the main body to complete the graduation thesis, and they should ensure sufficient time and energy to invest in professional research and thesis writing, so as to improve their ability to solve problems, further enhance their core competitiveness, and systematically present their professional knowledge through the thesis assessment. Secondly, in terms of guidance teachers, students should be encouraged to complete phased tasks in a timely manner, regularly check and help students solve problems in their papers, and provide timely guidance for knowledge in professional fields. Finally, in terms of university policies, universities should timely warn through the statistics of the graduation thesis management system and incorporate them into the evaluation and monitoring system, so as to manage the graduation thesis data system more effectively.

Acknowledgment

I received a lot of support throughout the writing process.

First of all, I would like to thank the support of the 2021 Industry-University Collaborative Education project of the Ministry of Education "Research on Graduation Thesis Process Management Strategy from the Perspective of Behavioral Data Analysis" (202101395039).

Then, I would like to thank VIP GRADUATION PROJECT MANAGEMENT SYSTEM for providing data, which is the data of students in 2021 and 2022 of an application-oriented undergraduate college in China.

References

- [1]. Zhang Zhe. Application Research of IT supplier evaluation based on AHP and Fuzzy Comprehensive evaluation [D]. Yunnan university of finance and economics, 2023. DOI: 10.27455 /, dc nki. Gycmc. 2023.000097.
- [2]. Wang Jianyu. Research on fuzzy comprehensive evaluation of environmental risk in drilling engineering of sulfur-bearing gas Wells based on AHP-entropy weight method [D]. Chongqing university, 2022. DOI: 10.27670 /, dc nki. Gcqdu. 2022.003480.
- [3]. Zhang Y. Application research based on fuzzy Hierarchical Analysis and fuzzy comprehensive evaluation model [D]. Anqing normal university, 2022. DOI: 10.27761 /, dc nki. Gaqsf. 2022.000014.
- [4]. Yu Chenxiong. Research on Public transport security risk evaluation based on fuzzy comprehensive evaluation method [D]. Northeast university of finance and economics, 2022. DOI: 10.27006 /, dc nki. Gdbcu. 2022.000907.
- [5]. Xu Jing. Application of Fuzzy Comprehensive Evaluation Method in quality evaluation of undergraduate graduation thesis [J]. Journal of tonghua normal university, 2019, 40 (12) : 18-21. DOI: 10.13877 / j.carol carroll nki cn22-1284.2019.12.005.
- [6]. Bei Jinlan, Wang Chunyu. Based on PDCA cycle improve undergraduate course graduation thesis quality [J]. Science and technology and enterprise, 2011 (16) : 219-220. The DOI: 10.13751 / j.carol carroll nki kjqy. 2011.16.144.

