Student Selection for Excellent Engineer Training: a Statistical Approach Based on Exam Scores

Linhua Sun
School of Earth Science and Engineering, Suzhou University, 49 Bianhe Road, Suzhou, Anhui 234000, China
E-mail of the corresponding author: sunlinh@126.com

Abstract
"Excellent engineer training program" is important for the education reform of China in recent years and a large number of studies related to the issue have been reported with lacking of discussion related to geological specialty. In this study, statistical analysis has been carried out based on the exam scores of geological students (sixty-four persons and twelve courses) in Suzhou University for revealing the differences between students and subdivision the students into different categories. The results of two-step cluster analysis suggest that these students can be subdivided into two groups with different characteristics, similar to the results obtained by factor analysis that twenty-nine students with high scores of English, global information system, structural geology, survey, palaeontology, coal geology, hydrology, geophysics and quaternary are considered to be chosen for training of excellent engineers.
Keywords: student selection, excellent engineer, statistical analysis, geological engineering

1. Introduction

"Excellent engineer training program" is an important reform of engineering education in universities and colleges in China at the moment. The purpose of this program is training a group of talented persons with innovative ability for adapting to economic and social development, which will be important for providing a solid pool of talent and intelligence to ensure for the construction of innovative country and, enhancing the country's core competitiveness (Lin, 2010a and b).

Because of the importance of this program, a series of studies have been paid towards to the training model, curriculum system and practice teaching methods for specialities of surveying and mapping, civil engineering, computer science and automation et al (Dong and Lu, 2011; Dong et al., 2011; Jiao and Zhou, 2012; Wu and Zhang, 2011; Zhang et al., 2012). However, similar work related to the geological speciality is lacking during the recent years.

"Excellent engineer training program" is a comprehensive research project including construction of teaching staff and experimental practice conditions, reform of training mode, as well as revision of mechanism related to students' selection and elimination. Among these procedures, students' selection plays an important role as students are the most basic raw materials for excellent engineer training.

In this paper, statistical methods, including two-step cluster and principle component analysis, have been applied for the exam scores of sixty-four geological students from Suzhou University, China, the goals of the study include: (1) classifying the students into different categories based on their examine scores and (2) identifying the differences between students. The study will be useful for future education because different students need different teaching methods.

2. Background

Suzhou University was founded in 1949, formerly known as a normal school. With the economic and social development, the school has grown into a science and engineering-based local universities. The specialty of geological engineering was established in 2007, and it has become a specialty closely associated with the development of local economic (coal mine) after five years of development. A total of 260 students are under-graduated and fifty-three have already graduated.

There are 20 full-time teachers with high levels of theory and practice of geology and engineering in the university, and most of them have been trained for teaching, scientific research and practicing about geology and coal mining. During the last five years, more than seventy papers have been published and nearly thirty of them have been indexed by SCI and Ei. Moreover, a large number of practice places have been established for the specialty education, including a basic geological training centre, an Engineering Research Centre of Coal Exploration and some other practicing bases. Experiments related to minerals, rocks, palaeontology, geological structure and microscope et al. can be processed in the centre, and some other advanced instruments, including plasma spectrometers, ion chromatography, X-ray fluorescence spectrometer, BET surface area analysis, seismographs, electrical, instrument, material testing machine and torsion testing machine etc. can be used for training the scientific abilities of the students. Moreover, the practicing bases in the coal mines around Suzhou City and Chaohu City were used for their practical ability training.

3. Methods

3.1 Data collection

All of the data for analysis are collected from the examination results during 2011, a total of sixty-four persons and twelve courses have been collected. The courses include english (EN), global information system (GIS), structural geology (ST), survey (SU), paleontology (PA), coal geology (CG), hydrology (HD), geophysics (GP), quaternary (QU), political (PO), physics (PE) and computer (PC). All of the scores are percentile.

3.2 Data analysis

There are two types of statistical analysis have been reported in this paper, the first one is two-step cluster analysis and another is factor analysis. All of them are processed by using SPSS (version 11). A detailed introduction about them can be found in the help topics of SPSS.

4. Results and discussions

4.1 Descriptions

The minimum, maximum and mean scores of the twelve courses of the sixty-four geological students from Suzhou University are given in Table 1 and shown in Figure 1. Standard derivation and skewness values are also shown in the table for identifying the degrees of score variations and the normal distribution of scores.

As can be seen from the table and figure, high mean scores are found for PC, PE, CG and HD with their values higher than 80, while low mean score is found for ST with a value lower than 70. Average score of PC is 88.1, followed by PE, CG and HD at 86.8, 85.4 and 80.5, respectively. The decreasing order of average scores is as follows: PC > PE > CG > HD > QU > EN > SU > PA > PO > GP > GIS > ST.

It can also be obtained from the standard derivation in Table 1 that the scores of PC have the minimal variation as its standard derivation is only 1.6, whereas the scores of ST have the largest variations as its standard derivation reached to 15.0. Additionally, scores of most of the courses except for PO, QU and PC are follow a normal distribution because the skewness values for PO, QU and PC are >= 1.0 or <-1.0.

4.2 Two-step cluster analysis

Two-step cluster analysis can use either schwarz's bayesian criterion (SBC) or akaike's information criterion (AIC). However, they produced same results and, therefore we report only the results of one type of analysis (SBC).

The results of two-step cluster analysis for the twelve courses of sixty-four students show that the students can be subdivided into two groups: the first group consists of twenty-five persons (39.1%) whereas the second group consists of thirty-nine persons (60.9%). Students in the first group are characterized by higher scores than the students in the second group.

The continuous variable-wise importance of course scores for group one and two are shown in Figure 2 and 3. As can be seen from the figure, the line of 95% confidence has crossed HD, PA, PO, CG, QU, GP, GIS, SU, ST and EN with positive student's t values. Comparatively, the line of 95% confidence has crossed SU, GP, CG, PA, GIS, HD and ST with negative student's t values. Such results suggest that most of the differences between these two groups are related to their specialized courses.

Table 1. Descriptive statistics of the scores (information about the abbreviations such as EN and GIS etc. see in the text)

Course	N	Min	Max	Mean	Sd	Skewness	Std. Err
EN	64	60	95	77.5	7.7	-0.3	0.3
GIS	64	61	93	74.1	8.3	0.1	0.3
SU	64	62	97	76.8	8.1	0.7	0.3
ST	64	27	91	63.1	15.0	-0.3	0.3
QU	64	16	97	79.2	14.6	-1.7	0.3
PA	64	62	91	76.0	8.4	0.0	0.3
PO	64	37	86	75.2	10.4	-2.3	0.3
PE	64	65	99	86.8	8.6	-0.7	0.3
CG	64	67	99	85.4	8.3	-0.1	0.3
HD	64	61	94	80.5	7.9	-0.2	0.3
GP	64	59	93	74.5	9.5	0.1	0.3
PC	64	86	93	88.1	1.6	1.0	0.3

Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.5, No.5, 2014

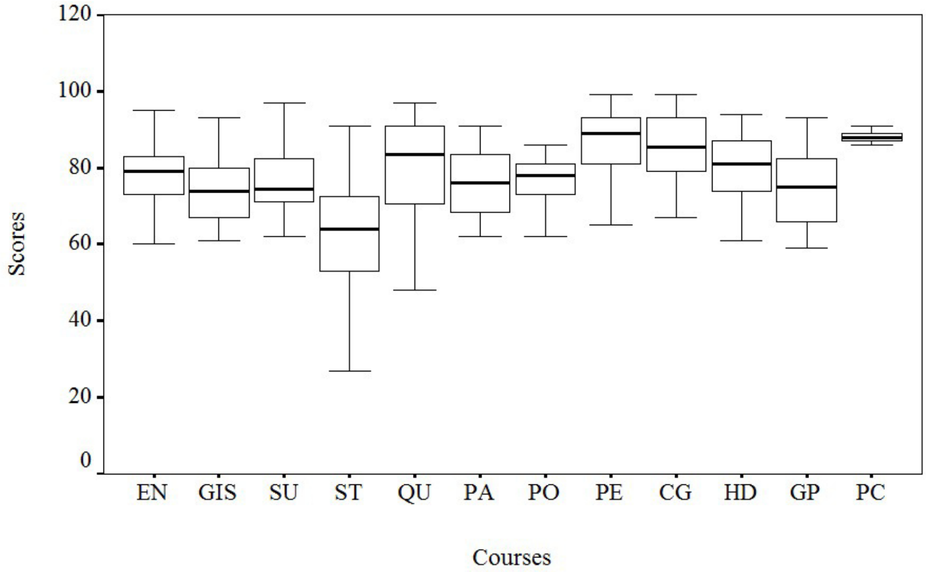

Figure 1. Boxplot of the scores

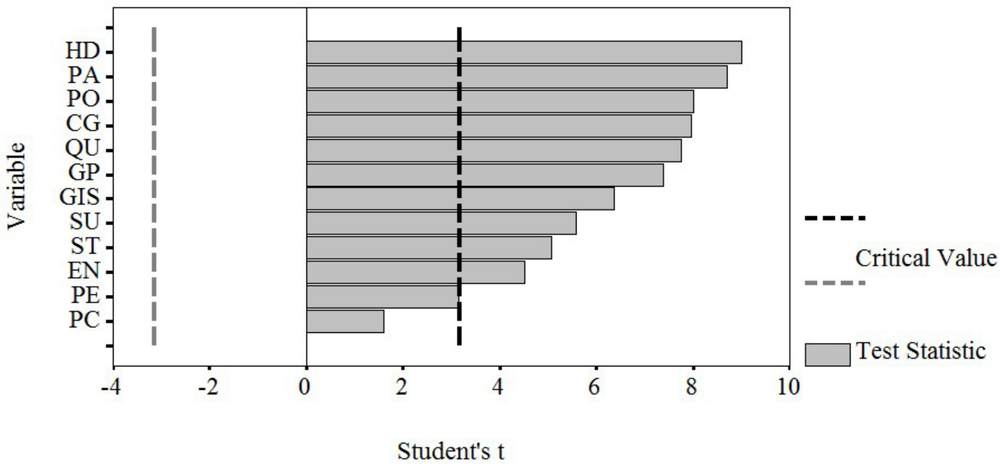

Figure 2. Bonferroni adjustment applied of group 1

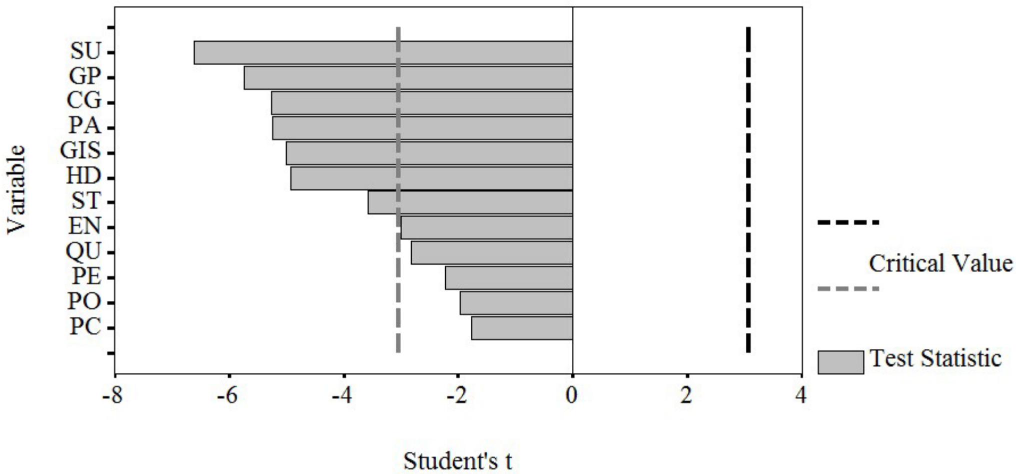

Figure 3. Bonferroni adjustment applied of group 2

4.3 Factor analysis

Factor analysis involves a mathematical procedure that transforms a number of correlated variables into a number of uncorrelated variables called principal components. The first factor accounts for as much of the variability in the data as possible, and each succeeding factor accounts for as much of the remaining variability as possible. The method has long been used for source identification of pollutants such as heavy metals (Bai et al., 2012).

The analytical results are listed in Table 2. As can be seen from the table, there are three factors have been obtained with eigen-value higher than one and a total variance explained is 70%. The first factor (FA1) accounts for 52.4% of the total variance and shows high positive loading on most of the courses, including EN, GIS, SU, ST, PA, CG, HD, GP, QU and PO. However, FA2 and 3 accounts for only 9.1% and 8.5% of the total variance, respectively, and there are two courses (PC and PE) show high loadings over them. The results indicate that most of the courses have similar effects for identifying the ability of the students.

For classification of the students by their exam scores, the factor score plots of FA1 versus FA2 (Figure 4) have been applied. The results suggest that the students can also be subdivided into two categories: the first group with twenty-nine students has positive FA1 scores, whereas group two students have negative FA1 scores. The results are similar to the subdivision by two-step cluster analysis.

Table 2. Results of factor analysis with eigen-value higher than one

Factors	% Var	EN	GIS	SU	ST	PA	CG	HD	GP	QU	PO	PE	PC
FA1	52.4	0.71	0.75	0.82	0.73	0.86	0.80	0.84	0.86	0.70	0.61	0.42	0.37
FA2	9.1	0.24	0.01	0.14	0.04	0.03	-0.25	-0.18	-0.11	0.01	0.07	-0.55	0.76
FA3	8.5	-0.05	0.27	0.05	0.13	0.09	-0.04	-0.16	0.04	-0.37	-0.57	0.53	0.39

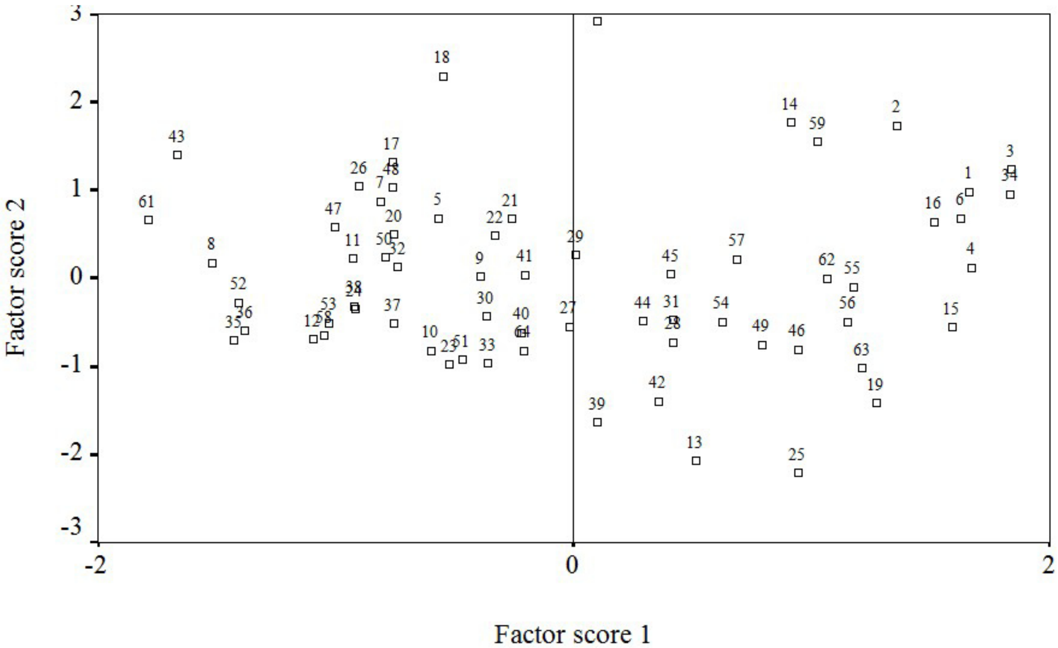

Figure 4. Factor score plots of cases (numbers are student IDs)

5. Conclusions

Scores of twelve courses for sixty-four geological students in Suzhou University have been collected for two-step cluster and principle component analysis, and the following conclusions have been made:

(1) The decreasing order of average scores of different courses is PC > PE > CG > HD > QU > EN > SU > PA > PO > GP > GIS > ST, and most of the scores show normal distribution except for PO, QU and PC;

(2) Two-step cluster analysis suggests that the students can be subdivided into two group (25 versus 39 persons) based on their specialized courses such as SU, GP, CG, PA, GIS, HD and ST, and it is similar to the results obtained by factor analysis that a classification of two group of students have been produced (29 versus 35) based on their SU, PA, CG, HD and GP scores;

(3) According to the results obtained by two-step cluster and factor analysis, twenty-nine students with high factor scores can be chosen for the training of excellent engineers.

Acknowledgement
This work is financed by Educational Quality Projects in Anhui Province, China. No. 20101060.

References

Bai, J. H., Xiao, R., Cui, B. S., Zhang, K. J., Wang, Q. G., Liu, X. H., Gao, H. F. & Huang, L. B. (2012). Assessment of heavy metal pollution in wetland soils from the young and old reclaimed regions in the Pearl River Estuary, South China. *Environmental Pollution*, 159(3), 817-824.

Dong, Q., Liu, D. Y. & Huang, L. Q. (2011). Practical teaching system of excellent civil engineers. *China University Teaching*, 1, 77-80.

Dong, X. P. & Lu, Z. Q. (2011). Curriculum system for excellent engineer training program of automation. *Research and Exploration in Laboratory*, 30(1), 265-267.

Jiao, M. L. & Zhou, L. (2012). The research of training mode of excellent engineer for surveying and mapping speciality. *Bulletin of Surveying and Mapping*, 1, 102-104.

Lin, J. (2010a). The development of general standards for "A plan for educating and training outstanding engineers. *Research in Higher Education of Engineering*, 4, 21-29.

Lin, J. (2010b). On the university working program of "A Plan for Educating and Training Outstanding Engineers". *Research in Higher Education of Engineering*, 5, 30-36.

Wu, S. Y. & Zhang, Y. H. (2011). Research and practice on the teaching model of outstanding engineers. *Research in Higher Education of Engineering*, 6, 34-39.

Zhang, L. X., Wang, Z. D., Wang, J. W., Hao, X. G. & Liu, S. B. (2012). Reform of experimental teaching system for excellent engineers training. *Education Research Monthly*, 1, 107-108.