The Comprehensive Quality Evaluation of Minority Students in Colleges and Universities Based on Principle of Information Entropy

Wu Wang[1], Lina Shan[2(✉)], and Yunjie Gu[2]

[1] Youth League Committee of Inner Mongolia University of Technology,
Huhehot 010051, Inner Mongolia Autonomous Region, China
[2] School of Reserve Officers, Harbin University of Science and Technology,
Harbin 150080, Heilongjiang, China
`linashan@l26.com`

Abstract. The paper does the analysis to the quality character of minority students in Colleges and Universities and there are four evaluation indexes which are suitable for the students: language and communication skills, academic performance and professional skills, personality accomplishment and psychological quality and democratic spirit and patriotism. Based on principle of information entropy, it analyzes the random sample of Inner Mongolia University of Technology graduate and makes use of the entropy to confirm the weight of each evaluation index to the influence of the comprehensive quality. It builds the linear evaluation model to get the comprehensive evaluation index and compares the result calculated by model and the student graduation test to prove the feasible of model with Pearson correlation coefficient.

Keywords: Minority students · Evaluation index
Principle of information entropy · Comprehensive quality

1 Suggest Problems and Confirm the Evaluation Index

The cultivation of the whole quality of the minority nationalities in West China affects the improvement of the western comprehensive quality [1]. Since it is in remote and developing area and most of minorities live in relatively independent region. They live in the cultural atmosphere of the nation, so their cognitive structure has obvious national character, which has an impact on the development of minority students in high school [2]. There are some common situations in the high school.

The normal education part of the minority students is finished in the local minority language before going to high school. The change of language after going into high school must cause difficulty in study and life. The students who are more adaptable to the language environment have obviously stronger ability to accept knowledge, compared with those students with poor language basis.

The basic courses of minority students are opened separately. In the teaching process, methods will be used mostly to finish teaching, such as reducing content, decrease the difficulty and reduce evaluation standard.

Because of the sudden change of culture environment atmosphere and the mutual communication with other nations, the students cannot adapt to the psychology, which has a great shock to their original knowledge to society and the sense of value.

The influence of minority students to different living custom and religion belief. They are very sensitive to nation, religion and belief and it is easy to have friction in communication. Whether they can communicate in high school harmoniously or not would affect their knowledge of national unity, the unity of the motherland.

In summary, the quality education of minority education have specialty. We cannot evaluate the minority students with the method of common high school students, but with these four aspects to have comprehensive evaluation: language and communication skills, academic performance and professional skills, personality accomplishment and psychological quality and democratic spirit and patriotism.

2 The Comprehensive Evaluation Model Based on Principle of Information Entropy

2.1 Information Entropy and Information Entropy Principle

Entropy is an important physical quantity. In the communication principle, the average information of source signal is called as entropy [3]. In information theory, the information entropy shows the degree of disorder of information [4]. The smaller the entropy is, the larger the information effect is; the larger the entropy is, the smaller the information effect is. The definition and formula of information entropy function is:

$$H(\vec{P}) = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad (1)$$

$\vec{P} = (p_1, p_2 \cdots p_n)$ stands for the probability vector.

2.2 Comprehensive Evaluation Model

We samples 50 minority students in Inner Mongolia University of Technology randomly and get the single assessment of four aspects (hundred-score system) language and communication skills, academic performance and professional skills, personality accomplishment and psychological quality and democratic spirit and patriotism by counselor and teacher evaluation, academic performance statistics, expert interview. The following picture is the score distribution histogram of 50 students.

Seen from the picture, these score distributions are independent and it has no crossing affect. The comprehensive quality evaluation model can be built with the data combined with principle of information entropy [5]. Following use n stands for student number ($n = 1, 2, \cdots \cdots 50$) and m stands for index ($m = 1, 2, 3 \ and \ 4$)

2.2.1 Confirmation of Information Entropy

Since all subjects score of the most of the sampling students is higher than 50, it is necessary to normalize the performance of each subject in each random sample to reduce the effect of the invalid information on the results. The specific operation is as follow: the result matrix is normalized and the normalized factor matrix is obtained.

$$z_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

The x_{max} stands for the highest of the single subject. The x_{min} stands for the lowest of the single subject.

According to the definition of entropy, the m subjects score of n students can be evaluation index and the entropy of evaluation index can be confirmed (Fig. 1).
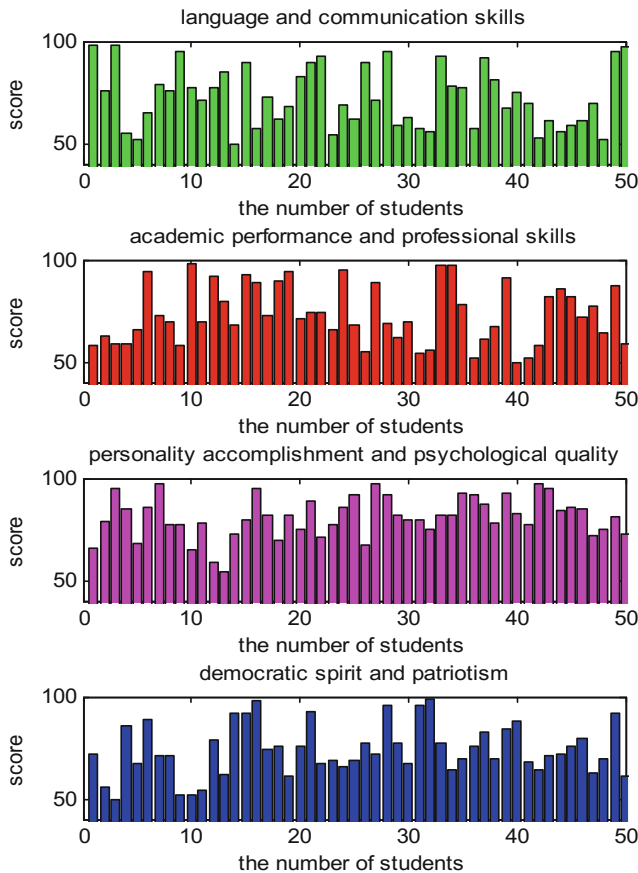

Fig. 1. The distribution of subjects of random sample students

$$H_j = -\frac{1}{\ln n} \sum_{i=1}^{n} (f_{ij} \ln f_{ij})$$

$$i = 1, 2, 3 \cdots n, j = 1, 2, 3 \ and \ 4$$

(3)

2.2.2   Revise the Traditional Probability Formula

The traditional probability formula is as follows.

$$f_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij}}$$

(4)

To make the H_j meaningful in the probability interval; we suppose that $f_{ij} = 0$. But since the logarithm of 1 is 0, so when $f_{ij} = 1$, $f_{ij} \ lnf_{ij}$ equals to zero, which is opposite to the definition of entropy. The reason why there is no such situation in traditional probability formula is that the probability of 1 of the signal is to determine the signal, which has no study significance. However, in this paper $f_{ij} = 1$ represents is a student whose score of subject j is the maximum value of random samples, which is obviously reasonable. Hence, the traditional calculation should not be adopted in the calculation. So the formula is revised as follows.

$$f_{ij} = \frac{1 + z_{ij}}{\sum_{i=1}^{n} 1 + z_{ij}}$$

(5)

2.2.3   The Calculation and the Nature of Weight

The principle of weight calculation should make fully use of the concept of information entropy. The smaller the entropy is, the larger the information effect is; the larger the entropy is, the smaller the information effect is. If score distribution presented is for the score of one subject, the entropy of the performance of 50 students will be smaller. Almost no one can get the same score, which means that the fractional division in the process of evaluation is detailed [6]. Almost every point has a distinction, so the objective factor is small, so the weight is larger. If the entropy of the performance of 50 students is larger, it means that the fractional division in the process of evaluation is rough and the subjectivity of the score is larger. So it will be affected by personal play and the subjective influence of the judger easily. So the model should be given less weight [7].

$$\omega_j = \frac{1 - H_j}{m - \sum_{j=1}^{m} H_j}$$

(6)

The nature of weight:

$$\sum_{j=1}^{m} \omega_j = 1 \tag{7}$$

We process with 50 random sample and induce the percentile data of four aspects which are language and communication skills ($j = 1$), academic performance and professional skills ($j = 2$), personality accomplishment and psychological quality ($j = 3$), democratic spirit and patriotism ($j = 4$) into MATLAB programming calculation weights [8]. The following are the result.

$$\begin{aligned} \omega_1 &= 0.3248, \ \omega_2 = 0.2970, \\ \omega_3 &= 0.1569, \ \omega_4 = 0.2213 \end{aligned} \tag{8}$$

2.2.4 Comprehensive Evaluation Model

According to the calculated weight, the comprehensive evaluation model can be described with four element linear equation. The comprehensive quality:

$$Q_i = \sum_{j=1}^{4} \omega_j \cdot x_{ij} \tag{9}$$

3 Comparative Argument

We can use the comprehensive quality model to calculate 50 random samples and compare the result gained from the model and the data gained form the graduation test. The results are shown as following picture (Fig. 2):

The relevance of data of two groups is counted with Pearson correlation coefficient. The Pearson correlation coefficient is also called as product difference correlation (product moment correlation) [9, 10]. It was a method to calculate linear correlation mentioned by the British statistician Pearson in 20 century. The correlation coefficient is larger, the correlation coefficient is more close to 1 and −1; the correlation coefficient is stronger, the correlation coefficient is close to 0, the correlation coefficient is weaker. The formula is as followed:

$$p_{xy} = \frac{N \sum X \cdot Y - \sum X \cdot \sum Y}{\sqrt{N \sum X^2 - (\sum X)^2} \cdot \sqrt{N \sum Y^2 - (\sum Y)^2}} \tag{10}$$

When deviation of the two variables is not zero, the correlation coefficient is defined. Pearson correlation coefficient is suitable for:
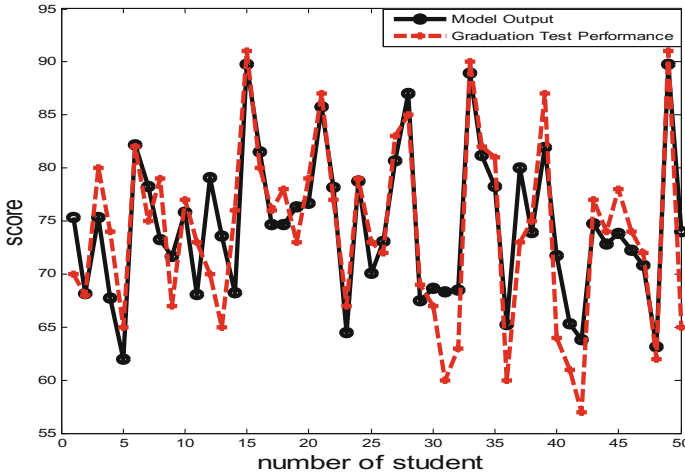
Fig. 2. The compare of graduation test performance and model output result

(1) A linear relationship between the two variables is the continuous data.
(2) The population of the two variables is the normal distribution, or close to the normal one.
(3) The observed values of the two variables are in pairs, and each pair of observations is independent of each other.

Suppose the graduation test result is X and the model calculation result is Y to count the Pearson correlation coefficient of the two by formula and the correlation coefficient is gained. Compare with the table of commonly used Pearson correlation coefficient (Table 1).

Table 1. Commonly used pearson correlation coefficient

Pearson correlation coefficient	Correlation degree
0.8–1.0	Extremely strong correlation
0.6–0.8	Strong correlation
0.4–0.6	Moderate correlation
0.2–0.4	Weak correlation
0–0.2	Extremely weak correlation

It is found that two parties is extremely strong correlation, which says that the model construction is successfully.

4   Conclusion

Clarify the evaluation index of minority students in high school and it is meaningful to build up reasonable evaluation model in the growth of the students and the mutual development of all nations. There are some conclusions gained from the model:

(1) Language communication skills have a lager effect to the development of the students in high school because it affects the improvement of all quality of the student directly. High school should think highly of the disadvantage of language communication and add some language training course.

(2) Besides the academic performance and professional skills, personality accomplishment and psychological quality, democratic spirit and patriotism take larger percentage in the comprehensive quality. High school should think highly of the ideological education to train the correct sense on life, world sense and value sense.

References

1. Wang, D.: POME-based fuzzy optimal evaluation model of water environment. J. Hehai Univ. **30**, 56–60 (2002)
2. Li, J.: The quality character of minority students in colleges and universities. High. Educ. Sci. **4**, 49–54 (2000)
3. Liang, J.Y., Shi, Z.Z.: The information entropy, rough entropy and knowledge granulation in rough set theory. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **12**, 37–46 (2004)
4. Liang, J., Shi, Z.D., Li, D., Wierman, M.J.: Information entropy, rough entropy and knowledge granulation in incomplete information systems. Int. J. Gen Syst **35**, 641–654 (2006)
5. Zhang, J.Y., Shu, L.X., Chen, S.X., Luan, X.X., Tian, T.: Study on the reserved model of information entropy for the calculation of the amount of nucleus radiant prevention drugs. J. Pharm. Pract. **19**, 765–777 (2008)
6. Huang, D.C.: Study of the comprehensive quality evaluation system for college students. Heilongjiang Res. High. Educ. **5**, 14 (2001)
7. Loucks, D.P., Gladwell, J.S.: Sustainability Criteria for Water Resource System. Tsinghua University Press, Beijing (2002)
8. Liu, Z.Y.: Scientific Computing and MATLAB. Science Press, Beijing (2001)
9. Yin, X.C., Liu, Y.Q.: Reforming experiment teaching to training students' comprehensive quality. Lab. Res. Explor. **2** (2003)
10. Zheng, R.Y, Lin, A.Z, Bao-Fu, L.U.: Reforming traditional mode of practice teaching and strengthening students' comprehensive quality. Res. Explor. Lab. **3** (2004)