



Research on influential factors of CET 4 results by using analysis methods

^{1*}Guanchen Zhou, ²Yalou Liu, and ³Yingna Zhao

¹Qingong College, Hebei United University, Tangshan, China

²College of Foreign Languages, Hebei United University, Tangshan, China

³Hebei United University, Tangshan, China

ABSTRACT

First, through analyzing the relation between the results of the subjects of the students and CET 4 pass rate, on the basis of the random sampling, the multiple correlation analysis has been implemented towards CET 4 passing results and other examination results of 25 non-English major students who pass CET 4 when attending CET 4 examination at the first time in one independent institute. The result shows that the results of CET 4 are not only related to the English learning ability, but also is connected with the thinking ability when the students are learning other subjects. Secondly, on the basis of the passing rate of CET 4, we use analytic hierarchy process (AHP) to analyze the English usual results, English learning time during spare time, English result of college entrance examination, teaching attitude, and gender and compare that which factor can provide the greatest impact on the results of CET 4. The research result shows that the importance of self English study. Finally, according to the different English learning basis and learning capabilities, we randomly choose 50 students in this independent institute and study the CET 4 results. We use the analysis of variance to study the average scores of listening, reading comprehension, writing and comprehensive test of male and female students respectively. The results of the above quantitative and qualitative analysis can help to improve English study capability in the future English learning process.

Keywords: CET 4; multiple linear regression; hierarchical analysis; analysis of variance

INTRODUCTION

CET is a national English examination. Since the first CET 4 in Sep. 1987, the college English examination has been spread for 26 years. Now it becomes the largest and most influential foreign language examination. Most colleges and universities think highly of CET 4, and CET 4 certificates would influence the graduation of the students. Because of the attention from the colleges and universities, the teachers and the students have to think highly of CET 4. English learning can be influenced by both the subjective factors of the learners, such as ages, study motivation, cognitive style and gender, and the objective factors, such as the English usual results, English learning time during spare time, English result of college entrance examination, and teaching attitude[1]. And the English usual results and the English learning time during spare time have the great impact. However, CET 4 is the decisive ruler to measure the English proficiency of the college students, and even some colleges and universities use it as one of the factors to influence the graduation of the students. So CET 4 and CET6 bring the negative effect on English teaching and learning. The result of the teachers group shows that 1/3 teachers say that their colleges have been developing the English teaching about CET 4, which is the same to the result of the students group. Over 30% students believe that the English teaching can not improve their English proficiency, but only is to the test[2]. Since Dec. 2013, CET has changed to pay more attention to the communication ability of English, which define a new target of English learning. Through the analysis of the influential factors of CET 4 results, we solve the problems of English learning and improve the English learning comprehensive ability of college students.

We have made a lot of reforms in CET 4. And this reform pays more attention to the practical application. Because the students haven't mastered the rules of the exam, although they spent a lot of energies on the preparation, they still couldn't get the satisfying results. But on the contrary, some students who spent less energies on it and pay more attention to the learning and thinking ability can get excellent results. Therefore, we further discuss the factors which can influence the results of CET 4.

2 Application of Multiple Linear Regression Model

In the regression analysis, if there are two or two more independent variables, it can be called multiple regression. Actually, one phenomenon may be related to many factors. When using the optimal combination of many independent variables to predict or estimate the independent variable is more effective and practical than just using one independent variable[3]. Therefore, it is more practical by using multiple linear regression than using simple linear regression.

After the analysis of a large amount of data, we know that the study ability of many subjects and the examination results may influence the CET 4 results. Because there a lot of influential factors, in this article we just suppose there are only 4 influential factors, such as the subjects of electronic technology experiment of the students, engineering mathematical experiment, circuit experiment and college English. Then the following multiple linear regression equation has been established, so that we can analyze the impacts of the subjects on the results of CET 4 accurately.

2.1 Suppose Random Variables

Set the result of CET 4 is variable y and the subjects of analog electronic technology experiment, engineering mathematical experiment, circuit experiment and college English are variables x_1, x_2, x_3 and x_4 . Suppose the relation of variable y and other four variables x_1, x_2, x_3 and x_4 are linear. The experimental data of the a th student is as below,

$$(y_a; x_{a1}, x_{a2}, x_{a3}, x_{a4}), \quad a = 1, 2, \dots, 25$$

Then the group of data can be listed below,

$$\begin{cases} y_1 = b_0 + b_1x_{11} + b_2x_{12} + \dots + b_4x_{15} + \varepsilon_1 \\ y_2 = b_0 + b_1x_{21} + b_2x_{22} + \dots + b_4x_{25} + \varepsilon_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_{25} = b_0 + b_2x_{251} + b_3x_{252} + \dots + b_4x_{255} + \varepsilon_{25} \end{cases}$$

$b_0, b_1, b_2, b_3, b_4, b_5$ are six estimated parameters; x_1, x_2, x_3, x_4 and x_5 are five common variables which can be accurately measured. $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{25}$ are 25 random variables which are independent and obey the same normal $N(0, \sigma)$. This is the mathematical model of multiple linear regression.

Suppose

$$Y = \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{25} \end{Bmatrix}, \quad X = \begin{Bmatrix} 1 & x_{11} & \dots & x_{14} \\ 1 & x_{21} & \dots & x_{24} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{251} & \dots & x_{254} \end{Bmatrix},$$

$$b = \begin{Bmatrix} b_0 \\ b_1 \\ \vdots \\ b_4 \end{Bmatrix}, \quad \varepsilon = \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{25} \end{Bmatrix},$$

Then the multiple linear regression model (2) can be written into matrix $Y = Xb + \varepsilon$. ε is 38-dimensional random vector and the sub-vectors are independent to each other.

Using software spss, we can obtain regression coefficients which are as follows,
 $b_0=157.037, b_1 = -1.295, b_2 = -0.507, b_3 = -0.436, b_4=4.735$.

The regression equation is $y = 157.037 - 1.295x_1 + 0.507x_2 - 1.295x_3 + 4.735x_4$.

According to the above obtained weight coefficients and regression equation, we can know that the impacts of the subjects on the results of CET 4 which is $x_4 > x_1 > x_2 > x_3$. This shows that the English usual result has greatest impact; the results of analog electronic technology experiment and engineering mathematics have less impact; and the result of circuit experiment has the least impact[4]. The results of engineering mathematical experiments and college English usual results are positive correlated to the result of CET 4. The analog electronic technology experiment and circuit experiments nearly can not improve the results of CET 4.

3 Application of AHP Model

The hierarchical analysis is to divide the overall target, sub-goals, evaluation criteria and the specific alternative solution into different hierarchical structure and then use the method of the judgment matrix eigenvector to get the optimal weights of each element of each layer to the element in the upper layer. Finally we use the method of weighted sum to hierarchically merge the ultimate weights of the alternative solutions to the overall goal. The greatest weight is the optimal plan. The optimal weight is a relative measure, which can show the evaluation criteria or sub-goal of each alternative solution of certain character[5]. It also can mark the relative measure of the superior degree and the importance relative measure of each sub-goal to the goal in the upper layer.

AHP is suitable for the goal system which has the layered staggered evaluation indicators. It can provide and analyze the quantitative basis through the comparison of various correlated factors layer by layer. In this article, we choose English usual results, time of English learning, English result of college entrance examination and teaching attitude to implement the comprehensive evaluation[6]. We also provide the weights of these factors and divide these problems into two layers. The upper layer, represented by A , is goal layer which is the comprehensive evaluation of CET 4 results. The second layer, represented by B_i , is the factor layer which is the English usual results. The impact of five factors B_1, B_2, B_3, B_4, B_5 in the comparison layer on the upper layer A can be shown in the following matrix.

$$A = (a_{ij})_{n \times n}, a_{ij} > 0, a_{ji} = \frac{1}{a_{ij}}$$

This kind of matrix can be called comparison matrix. In the matrix, factor a_{ij} represents the ratio of importance of B_i and B_j to A . According to the characters of the pairwise comparison matrix, we know $a_{ii} = 1$. If we construct a n order matrix, we just implement $\frac{1}{2}n(n-1)$ comparisons.

In order to calculate conveniently, the different scales rating is established as below,

Table 1 is the relative weight of each factor.

Scale a_{ij}	Meaning
1	B_i and B_j have equal importance.
3	The importance of B_i is little stronger than B_j .
5	The importance of B_i is stronger than B_j .
7	The importance of B_i is obviously stronger than B_j .
9	The importance of B_i is absolutely stronger than B_j .
2,4,6,8	The importance of B_i and B_j is in the between two adjacent levels.
1,1/2, ..., 1/9	The importance of B_i and B_j is opposite to the above mentioned.

Table 1 Relative weight of each factor

	English usual result	English learning time	English result of college entrance examination	Teaching attitude	Gender
English usual result	1	2	$\frac{3}{2}$	$\frac{3}{4}$	3
English learning time	$\frac{1}{2}$	1	$\frac{3}{8}$	$\frac{7}{9}$	$\frac{6}{5}$
English result of college entrance examination	$\frac{2}{3}$	$\frac{8}{3}$	1	$\frac{1}{2}$	$\frac{7}{3}$
Teaching attitude	$\frac{4}{3}$	$\frac{9}{7}$	2	1	$\frac{5}{4}$
Gender	$\frac{1}{3}$	$\frac{5}{6}$	$\frac{3}{7}$	$\frac{4}{5}$	1

The matrix is

$$A = \begin{bmatrix} 1 & 2 & \frac{3}{2} & \frac{3}{4} & 3 \\ \frac{1}{2} & 1 & \frac{3}{8} & \frac{7}{9} & \frac{6}{5} \\ \frac{2}{3} & \frac{8}{3} & 1 & \frac{1}{2} & \frac{3}{7} \\ \frac{4}{3} & \frac{9}{7} & 2 & 1 & \frac{5}{4} \\ \frac{1}{3} & \frac{5}{6} & \frac{3}{7} & \frac{4}{5} & 1 \end{bmatrix}$$

After we do the consistency test to A , we obtain the maximum eigenvalue $\lambda = 5.2608$ of A . Then we calculate the disposable indicator

$$CI = \frac{\lambda - n}{n - 1} = 0.0652$$

Because $CR = \frac{CI}{RI} = \frac{0.0652}{1.12} = 0.0582 < 0$, we can determine this matrix pass the consistency test, which means that the distribution of weights is reasonable. When $\lambda = 5.2608$, the corresponding eigenvector is $[0.2685, 0.1319, 0.2199, 0.2617, 0.1180]$.

So we know that the weight of English usual result is maximal, which means that among these factors, English usual result has the greatest impact. If the English usual result of the student is outstanding, the student has the greatest possibility to pass CET 4.

4 Research on the Difference of CET 4 Results Based on the Gender Difference By the Analysis of Variance

The analysis can testify the significance of the difference of the averages and also can judge the degree of the importance. The analysis of variance in the principle of statistics mainly includes univariate analysis of variance and multivariate analysis of variance. For one studied indicator X , it always influenced by some factors $A_1, A_2, A_3 \cdots A_n$, and it also will be influenced by the random deviation. We need to study how $A_1, A_2, A_3 \cdots A_n$ influence X and how the random deviation influences X . In this article we mainly discuss the analysis of variance of two factors.

We select 50 students from non-English major students in our college. Then we use the method of variance analysis

to analyze gender and the scores of each part of CET 4 results such as listening, reading comprehension, writing, and comprehensive test. Because we should consider the influence of random factors on the indicators, we suppose that all the data are from the same normal population. We study A (male) and B (female) equally. A has four different levels A_1, A_2, A_3, A_4 (scores of listening, reading comprehension, writing, and comprehensive test for male); B has four different levels B_1, B_2, B_3, B_4 (scores of listening, reading comprehension, writing, and comprehensive test for female). There is no interaction between A and B . We implement an independent test to the horizontal group ($A_i B_j$) and obtain ab experiment results $X_{ij} (i = 1, 2, \dots, a, j = 1, 2, \dots, b)$. And the experiment results are shown in Table 2.

Table 2 Two-factor no repeat test table

Factors	B_1	B_2	B_j	B_b	Average \bar{X}_{i-}
A_1	X_{11}	X_{12}	X_{1j}	X_{1b}	\bar{X}_{1-}
A_2	X_{21}	X_{22}	X_{2j}	X_{2b}	\bar{X}_{2-}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_i	X_{i1}	X_{i2}	X_{ij}	X_{ib}	\bar{X}_{i-}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_a	\bar{X}_{a1}	\bar{X}_{a2}	X_{aj}	X_{ab}	\bar{X}_{a-}

$$\text{And } X_{-j} = \frac{1}{a} \sum_{i=1}^a X_{ij} \quad (j = 1, 2, \dots, b) \tag{1}$$

$$X_{i-} = \frac{1}{b} \sum_{j=1}^b X_{ij} \quad (i = 1, 2, \dots, a) \tag{2}$$

Suppose X_{ij} is one selected sample from the normal distribution $X_{ij} \sim N(u_{ij}, \sigma^2)$. We suppose there is no interaction in A, B . Suppose $X_{ij} = u_{ij} + \varepsilon_{ij}$ and $\varepsilon_{ij} \sim N(0, \sigma^2) (i = 1, 2, \dots, a; j = 1, 2, \dots, b)$; u_{ij} is the theoretical expectation in the condition of $A_i B_j$; ε_{ij} represents the random deviation and the deviations are independent. According to (1), we can get

$$u = \frac{1}{ab} \sum_{i=1}^a \sum_{j=1}^b u_{ij} \tag{3}$$

$$u_{j-} = \frac{1}{a} \sum_{i=1}^a u_{ij} (j = 1, 2, \dots, b) \quad u_{i-} = \frac{1}{b} \sum_{j=1}^b u_{ij} (i = 1, 2, \dots, a)$$

Suppose $\alpha_i = u_i - u$ and $\beta_j = u_{-j} - u$; α_i is the i^{th} horizontal effect of factor A_i ; β_j is the j^{th} horizontal effect of factor B_j ; they represent the impact degrees of all the levels of A and B . The relation formula is

$$\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = 0 \tag{4}$$

Suppose $\delta_{ij} = u_{ij} - u_i - u_j + u$. There is interaction between A_i and B_j . Suppose there is no interaction

between A_i and B_j ; and $u_{ij} = u + \alpha_i + \beta_j$.

Then we can obtain the mathematical model of two-factor no repeat test variance.

$$\left\{ \begin{array}{l} X_{ij} = u + \alpha_i + \beta_j + \varepsilon_{ij}, i = 1, 2, \dots, b, \\ \sum_{i=1}^a \alpha_i = 0, \sum_{j=1}^b \beta_j = 0 \\ \varepsilon_{ij} \sim N(0, \sigma^2) \end{array} \right.$$

$$u, \sigma^2, \alpha, \beta \quad (j = 1, 2, \dots, a, j = 1, 2, \dots, b),$$

In Table 3 we select the scores of each part of CET 4 results (listening, reading comprehension, writing and comprehensive test) from 50 male and female students. After the analysis of the data in the table, we use SPSS to implement the multivariate analysis of variance and obtain Table 4 and Table 5.

Table 3 CET 4 results

Gender	Listening	Reading Comprehension	Writing	Comprehensive Test
Male	102	128	70	45
Male	128	126	86	30
Female	159	133	84	47
Female	123	125	67	40
...
女	107	154	75	30
女	117	152	76	40

Table 4 Descriptive statistics table

Descriptive Statistics				
Dependent variable: Scores of Each Part				
Gender	Parts	Average	Standard deviation	N
Male	Listening	117.24	17.683	25
	Reading Comprehension	135.48	22.070	25
	Writing	75.48	11.762	25
	Comprehensive Test	38.08	7.297	25
	Total	91.57	41.006	100
Female	Listening	117.96	21.289	25
	Reading Comprehension	138.04	18.533	25
	Writing	75.28	11.788	25
	Comprehensive Test	35.12	8.922	25
	Total	91.60	42.884	100
Total	Listening	117.60	19.372	50
	Reading Comprehension	136.76	20.211	50
	Writing	75.38	11.655	50
	Comprehensive Test	36.60	8.204	50
	Total	91.58	41.850	200

Table 4 is descriptive statistics. For males, their comprehensive test results are higher than those of the females. For females, the scores of reading comprehensive are higher. The scores of writing and listening parts are basically the same. From above, we can't judge that the score difference is caused by the random selection or by the gender difference. So we should analyze the other results of the analysis of variance.

Table 5 Levene of deviation variance equivalence

<i>Levene</i> Test * of Deviation Variance Equivalence			
Dependent Variable: Scores of Each Part			
F	df1	df2	Sig.
5.069	7	192	.121
In each group of dependent variables, the deviation variances are equal.			

<i>Levene</i> Test * of Deviation Variance Equivalence			
Dependent Variable: Scores of Each Part			
F	df1	df2	Sig.
5.069	7	192	.121
In each group of dependent variables, the deviation variances are equal.			
a. Design: Intercept + Gender + Parts of CET 4 + Gender * Parts of CET 4			

P in *Levene* test shown in Table 5, is 0.121 which is larger than 0.05. We can say that it satisfies the homogeneity of variance and pass the test. So we can draw the conclusion that in CET 4, male students are good at comprehensive test and female students are good at reading comprehension.

CONCLUSION

From the above statistic data, the results of CET 4 are related to the results of other subjects in a sense. The different thinking ability and study ability relate the results differently. The electronic technology experiment, engineering mathematical experiment, circuit experiment and college English have significant relation to the English learning. This shows that the increase of thinking ability can improve the English examination result.

Through the analysis of the above four subjects and the AHP model, we draw the conclusions. The English usual result has the maximum impact on the CET 4 results; the English usual results can directly reflect the English study ability of the students. Besides, other subjects can influence the results of CET 4 in a certain degree. So the thinking ability and study ability of other subjects can influence the English learning ability indirectly. We should cultivate the English learning ability from many aspects, such as the comprehension ability of the student, logic thinking ability, memory capability, abstract thinking ability, spatial imagination and comprehensive ability. So the assessment of English learning ability is the assessment of the comprehensive learning ability of the students. Only obtaining strong comprehensive quality, can the students get good results in CET 4.

From the analysis of variance model, we can get the following conclusion. In CET 4, male students are more outstanding than the female in the comprehensive test, while the female students are more outstanding than the male in reading comprehension. According to the conclusion, the students should make progress in their weak parts as well as maintain their advantages. In the future, the students should study actively, accumulate the knowledge, and cultivate their thinking style, study methods, logic thinking ability and social practice ability. On the basis of the present learning methods, they should develop their abstract thinking ability and logic thinking ability. They also should pay more attention to the exchange of learning methods and experiences so as to make improvement together.

Acknowledgements

We thank anonymous reviewers for helpful comments. This research is partially supported by the National Science and Technology Program Innovative Approaches Special Project "Application and Practice of Scientific Methods and Scientific thinking in University Teaching Innovation" (NO.2009IM010400)

REFERENCES

- [1] Xiang Xiaolin, Wang Wei, *Chongqing University Journal*, **2003**(9)4:71-74.
- [2] Xu Zhimin, Liu Xiyu, *Computer Engineering*, **2004**(30) 21: 129-131.
- [3] Anthong Woods, Paul Fletcher, Arthur Hughes. *Statistics in Language Study*. Beijing: Foreign Language Teaching and Research Press, UK: Cambridge Press, **2000**.
- [4] Xu Jun, Liu Na, Basic Idea and Practical Application of AHP, *Intelligence Explore*, **2008**,12:113-115.
- [5] Zhu Zhimin, *Qiqihar Teachers College Journal*, **2013**,3:143-145.
- [6] Han Baocheng, Dai Manchun, Yang Lifang, *Foreign Language Teaching and Research*, **2004**,2:17-23