Analysis on the Weights of Teaching Assesses by Using Multivariate Statistical Methods

Yuankuan Zhou*, Qin Xu, Wanjun Zhang
College of Educational Science, Anhui Normal University, Wuhu 241000, China
*Corresponding author(E-mail: zhouyuankuan@126.com)

Abstract
This paper discusses the role of multivariate statistical methods in determining the weights of each index in the teaching evaluation system and applies it to the discussion of the weights of teaching assesses of a university. First, we use the Likert scale to quantify the number of each option into the corresponding option score. Then we use the principal component analysis and the factor analysis method to develop the weights of each index, and get the evaluation score table of the university under different evaluation function. The wilcoxon signed-rank test compares the methods and finds that the multivariate statistical method is better than the old methods of determining the weights of the university. Therefore, the paper chooses the result of factor analysis as the evaluation function according to the practical significance.

Key words: principal component analysis, factor analysis, weights

1. INTRODUCTION

From the perspective of the world, after decades of development, student evaluation has become an important system of teaching quality assurance. To carry out student assessment work has opened a conventional channel for students to express their teaching minds, so that universities can respect students teaching willing and protect their dominant position. Students’ evaluation can help teachers and teaching managers understand the teaching situation, sum up their experiences to promote the professional development, improve the quality of teaching, and then help colleges and universities to adjust the school orientation, and establish the purpose of serving the students. However, the students' assessment of teaching brings improvement of the school's related work, it also exists some problems when practices it. As an important basis for evaluating the quality of teachers' teaching, the implementation of student evaluation system has aroused the greater response of teachers' groups, especially the questioning and worrying about the evaluation index system itself and its interpretation and application. In this regard, the scientific evaluation and reasonable set of student evaluation index system in the weight of the indicators, is a worthy of in-depth study of the problem.

From the actual situation, the evaluation method of each university is not the same. This paper chooses the evaluation data of the teaching effect of a university in Wuhu city of Anhui province as an example, and makes an empirical analysis on the effect of the evaluation method. The school's assessment methods include the following eight indicators: \( x_1 \): whether in the teaching process is seriously responsible, disciplined, pay attention to the teacher appearance; \( x_2 \): whether teaching preparation is full, lectures are proficient, serious and full of passion; \( x_3 \): teaching is clear, focused, linguistic and express accurately; \( x_4 \): good at inspiring induction, teaching methods are flexible and diverse; \( x_5 \): focus, difficulty, and depth are appropriate, put theory with practice; \( x_6 \): teaching content is full, novel and efficient; \( x_7 \): the basic concept is accurate and clear, logical structure is reasonable, the point of view is correct, clarity, systematic; \( x_8 \): students assessment of the whole teaching. In the evaluation process, the weights of each indicator are given directly related to whether the results of the evaluation are prepared, so this is a question worthy of further discussion. In this paper, the weights of the indicators in the teaching effectiveness evaluation methods are: 0.3, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1 The appropriate weighting method is given, and the appropriate weighting method is given according to the multivariate statistical theory. Finally, the wilcoxon signed-rank test is used to compare the advantages and disadvantages of the method and the original method.

2. STATISTICAL MODEL

The evaluation function of teaching quality effect is \( c(x) \), It is a linear function of \( x = (x_1, x_2, \ldots, x_8) \), which is
\[ c(x) = \sum_{i=1}^{k} \lambda_i x_i, \quad (1) \]

where \( \sum_{i=1}^{k} \lambda_i = 1, \lambda_i \geq 0, \quad i = 1, 2, \ldots, k. \)

we call \( \lambda_i \) is the weights of \( x_i \). When the students score every \( x_i \), it is possible to obtain a comprehensive value of the student's evaluation of the teaching quality of the teacher according to the evaluation function \( c(x) \).

The following is a multivariate statistical method to establish a model to determine the weights of each indicator in the evaluation function.

### 2.1. The Likert Scale

Before establishing the weight model to determine the indicators in the evaluation function, the first problem to be solved is how to quantify the evaluation of each index in the process of student evaluation. This paper uses the Likert scale to quantify the indicators. The Likert scale is a psychological response scale, which is often used in questionnaires and is the most widely used scale in current research. When the subjects respond to the items of such questionnaires, they specifically indicate their degree of recognition of the statement. It was made from the original total scale by the American social psychologist Li Kete in 1932.

According to the sample data of the quality evaluation of teachers in the university, the evaluation system contains 8 indicators, using the 5-level Likert scale, contains five kinds of answers: very satisfied, satisfied, basically satisfied, dissatisfied, very dissatisfied and they are recorded as 5, 4, 3, 2, 1 points as the weight, multiplied by the corresponding number of options, get a corresponding evaluation of the corresponding points.

### 2.2. Principal component analyses

Principal Component Analysis is a way to reduce the number of variables into a few variables by reducing the dimension of the method, but also a multi-index into a few comprehensive indicators of statistical methods. It was first proposed by Pearson in 1901, and Hotelling was developed in 1933. This method can summarize most of the information from the selected index system, synthesize it according to the information provided by the main component, and avoid the comprehensive evaluation the subjective influence. In the case of the least loss of data information, the original multiple indexes are translated into one or more comprehensive indexes by reducing the dimension, and the objective weighting is carried out according to the relative importance of the indexes. The specific modeling and calculation process is as follows:

1. **Construct the sample array**
   - The number of indicators for teaching quality evaluation is \( k \), the number of teachers to be rated is \( n \), the sample matrix is \( X \), \( x_{ij} \) indicates the \( i \) teacher's \( j \) indicators. Then we have:
   
   \[
   X = \begin{bmatrix}
   x_{i1} & \cdots & x_{ik} \\
   \vdots & \ddots & \vdots \\
   x_{ni} & \cdots & x_{nk}
   \end{bmatrix}
   \]

   where \( x_{ij} \) represents the value of the \( j \) variable in the \( i \) group sample data.

2. **Normalize the raw data**
   - Where \( z_{ij} = \frac{y_{ij} - \bar{y}_j}{s_j} \), \( \bar{y}_j \), \( s_j \) are respective the mean and the standard deviation of the \( j \) row in the X array.

3. **Calculate the correlation coefficient matrix**
   - \( R = \frac{Z^T Z}{n-1} \)

4. **Find the eigenvalues and eigenvectors of the correlation coefficient matrix \( R \).** Suppose \( R \) has a
characteristic root of \( \lambda_1 \geq \cdots \geq \lambda_k \geq 0 \), and the corresponding eigenvector is \( T_1, T_2, \ldots, T_k \), where \( T_{(m)} \) and \( T_{(k-m)} \) are the first \( m \) and the last \( k-m \) eigenvectors of \( R \) respectively.

(5) Take a linear combination of the top \( m \) principal components:

\[
\sum_{s=1}^{m} \lambda_s \gamma_s(x) = \hat{\lambda}_{(m)}' y(m) = \hat{\lambda}_{(m)}' U_{(m)}' x
\]

The right side of the equation \( U_{(m)} \hat{\lambda}_{(m)} \) is normalized so that it satisfies and is 1 after the resulting weight, resulting in an evaluation function.

2.3. Factor analysis

Factor analysis is also a multivariate statistical analysis method for dimensionality reduction. It is actually a generalization and development of the principal component analysis method used to analyze the factors that are hidden behind the surface phenomena.

In the factor analysis model \( X = AF + \varepsilon \), where \( X \) is the observable random vector, \( F \) is the common factor vector, \( \varepsilon \) is the special factor vector, \( A = (a_{ij})_{p \times m} \) is the factor load matrix. If we do not consider the influence of the special factor, when \( m = p \) and are reversible, we can easily calculate the corresponding value on factor \( F \) from the value of each sample's index \( X \), \( F = A^{-1}X \) the "score" of the sample on factor \( F \), referred to simply as the factor score for the sample.

However, the factor analysis model requires \( m < p \) in practical application. Therefore, it cannot accurately calculate the score of the factor, and only the factor score can be estimated. There are many ways to estimate the factor scores, the following Tom's return to the law .

The method assumes that the common factor can be regressed to \( p \) primitive variables,

\[
\hat{F}_j = b_{j0} + b_{j1}X_1 + \cdots + b_{jp}X_p, \quad j = 1, \ldots, m
\]

If \( F_j \) and \( X_i \) are normalized, the regression constant is zero, is \( b_{j0} = 0 \).

By the meaning of the factor load, we can see for any of \( i = 1, \ldots, p; \quad j = 1, \ldots, m \), we have

\[
a_{ij} = r_{ij} = E(X_iF_j) = E[X_i(b_{j1}X_1 + \cdots + b_{jp}X_p)] = b_{j1}E(X_iX_1) + \cdots + b_{jp}E(X_iX_p) = b_{j1}r_{11} + \cdots + b_{jp}r_{pp}
\]

If

\[
B = \begin{bmatrix}
 b_1 \\
 \vdots \\
 b_m
\end{bmatrix} = \begin{bmatrix}
 b_{11} & b_{12} & \cdots & b_{1p} \\
 b_{21} & b_{22} & \cdots & b_{2p} \\
 \vdots & \vdots & \ddots & \vdots \\
 b_{m1} & b_{m2} & \cdots & b_{mp}
\end{bmatrix},
\]

Then the above formula can be written in matrix form

\[
A = RB', \quad \text{or} \quad B = A'R^{-1}
\]

Such that

\[
\hat{F} = \begin{bmatrix}
 \hat{F}_1' \\
 \vdots \\
 \hat{F}_m'
\end{bmatrix} = \begin{bmatrix}
 b_{11}' & X \\
 \vdots & \vdots \\
 b_{m1}' & X
\end{bmatrix} = BX = A'R^{-1}X
\]

where \( m \) is the number of principal factors.

The variance contribution rate of all common factors \( \hat{F}_m \) is:

\[
v_m = \sum_{i=1}^{p} a_{im}^2
\]

Then the synthetic index is:

\[
y(x) = \sum_{i=1}^{m} v_i \hat{F}_i = \tau_1 x_1 + \tau_2 x_2 + \cdots + \tau_p x_p
\]
3. EMPIRICAL ANALYSIS AND RESULTS

3.1. Model solution

In order to determine the relationship between the extracted principal components and the original variables, we can use SPSS software to calculate the relationships among variables. From the calculation results, we can see that the correlation between the KMO=0.955. So variables are very strong, and the principal component analysis can be carried out. Next the sample data from the survey were standardizing handled, and the eigenvalues of the correlation matrix and the contribution rate of each principal component are also solved.

Starting with standardized data, first, the principal components of these indexes are calculated, and then sort them according to the size of the principal component. The results are listed in Tables 1 and 2. Table 1 is the information about the characteristic root, the variance contribution rate and the accumulative contribution rate, and Table 2 is the information about the eight vectors.

<table>
<thead>
<tr>
<th>Serial number of each index</th>
<th>Characteristic root value</th>
<th>Variance contribution rate (%)</th>
<th>Cumulative contribution rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.555</td>
<td>81.937</td>
<td>81.937</td>
</tr>
<tr>
<td>2</td>
<td>0.345</td>
<td>4.316</td>
<td>86.253</td>
</tr>
<tr>
<td>3</td>
<td>0.229</td>
<td>2.864</td>
<td>89.117</td>
</tr>
<tr>
<td>4</td>
<td>0.225</td>
<td>2.817</td>
<td>91.934</td>
</tr>
<tr>
<td>5</td>
<td>0.191</td>
<td>2.391</td>
<td>94.325</td>
</tr>
<tr>
<td>6</td>
<td>0.162</td>
<td>2.026</td>
<td>96.350</td>
</tr>
<tr>
<td>7</td>
<td>0.158</td>
<td>1.976</td>
<td>98.326</td>
</tr>
<tr>
<td>8</td>
<td>0.134</td>
<td>1.674</td>
<td>100.000</td>
</tr>
</tbody>
</table>

The result of model calculation shows that:

Using the first principal component synthesis (contribution rate of 81.937%), the evaluation function is:
\[ c(x) = 0.1228x_1 + 0.1251x_2 + 0.1265x_3 + 0.1255x_4 + 0.1245x_5 + 0.1250x_6 + 0.1255x_7 + 0.1252x_8 \]

Using second principal component synthesis (contribution rate of 86.253%), the evaluation function is:
\[ c(x) = 0.1341x_1 + 0.1300x_2 + 0.1318x_3 + 0.1171x_4 + 0.1199x_5 + 0.1166x_6 + 0.1262x_7 + 0.1243x_8 \]

Using third principal component synthesis (contribution rate of 89.117%), the evaluation function is:
\[ c(x) = 0.1331x_1 + 0.1367x_2 + 0.1279x_3 + 0.1188x_4 + 0.1181x_5 + 0.1145x_6 + 0.1199x_7 + 0.1309x_8 \]

3.2. Factor analysis

To compare with the above three evaluation functions, and then we carry on factor analysis. Here, in order to get the meaning factor of significance, we rotate the factor load matrix \( A^T \) to the maximum variance method. The factor loading matrix is obtained after rotation.

<table>
<thead>
<tr>
<th>Contribution rate</th>
<th>Cumulative contribution rate</th>
<th>Rotated Component Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.373</td>
<td>0.373</td>
<td>-0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.490</td>
</tr>
<tr>
<td>0.336</td>
<td>0.709</td>
<td>-0.121</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.584</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.459</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.276</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.643</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.499</td>
</tr>
</tbody>
</table>
Next, a comprehensive evaluation function of factor scores is given:

\[ c(x) = 0.1234x_1 + 0.1254x_2 + 0.1267x_3 + 0.1250x_4 + 0.1242x_5 + 0.1245x_6 + 0.1254x_7 + 0.1253x_8 \]

### 3.3. Consistency test

In the original method of the university, the weight of the first index is 0.3, and the other indexes are weighted 0.1, in the following tables, the comprehensive scores of teaching quality evaluation results are compared with other methods in the paper. Through the Wilcoxon signed-rank test, the goodness of fit values is shown in table 4.

<table>
<thead>
<tr>
<th>P</th>
<th>1 principal components</th>
<th>2 principal components</th>
<th>3 principal components</th>
<th>factor analysis</th>
<th>Original method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 principal components</td>
<td>1</td>
<td>0.6848</td>
<td>0.7395</td>
<td>0.9587</td>
<td>0</td>
</tr>
<tr>
<td>2 principal components</td>
<td></td>
<td>1</td>
<td>0.9286</td>
<td>0.7008</td>
<td>0</td>
</tr>
<tr>
<td>3 principal components</td>
<td></td>
<td></td>
<td>1</td>
<td>0.7548</td>
<td>0</td>
</tr>
<tr>
<td>factor analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From Table 4, it can be seen that the \( p \) value of the original method and other methods are far less than 0.05. That is, there is a clear difference. The comprehensive evaluation results of factor analysis are compared with those obtained by principal component analysis. The \( P \) values are greater. It shows that the results of factor analysis and other principal component analysis have the least difference among the methods; and from the significance of each index, the main factor extracted from factor analysis is more convincing. The first main factor 4,5,6 is larger, reflecting the teacher's teaching methods are recognized by students; the second main factor 1,3,7 is larger, reflecting the teaching attitude of teachers, clear thinking, logical structure; the third main factor 2,8 is larger, reflecting the students' overall evaluation of teachers' teaching. From the empirical analysis of the paper, it can be seen that the original method of the university has given too many weights to the index \( x_1 \), so it is necessary to improve the weight of each index in the comprehensive evaluation.

It is important to point out that the 8 evaluations provided by the sample show little difference and that the students are very small in scoring. Because principal component analysis and factor analysis are based on the score of students, the weight coefficients extracted from the common factor are also very small.

### 4. CONCLUSION

The paper discusses the role of multivariate statistical methods in determining the weights of each index in the teaching evaluation system and applies it to the discussion of the weights of teaching assesses of a university. We compares the method of determining the weight in the paper with the method of the university and find the multivariate statistical method is better than the old methods of determining the weights of the university. The result of the method of the university is obviously different from that of the method proposed in this paper. The original method of the university has given too many weights to the index \( x_1 \), so it is necessary to improve the weight of each index in the comprehensive evaluation. The view can also be explained in the light of pedagogy. The main representative of index \( x_i \) is the teaching attitude and it is a general basic requirement for teaching. Its specific connotation is obscure and the subjectivity in practice is strong. Then it will cause differences in evaluation among different students.

The weight coefficients in the paper are similarity each other, extracted from the common factor are also very small. This may be because the students have little difference in the score of the 8 evaluation indicators for the sample. Because principal component analysis and factor analysis are based on the score of students, the weight coefficients extracted from the common factor are also very small.
Reference


