

A Comprehensive Evaluation of the Mental Health Status of Military Personnel Based on Factor Analysis

Chunmei Chen*, Ping Yang, Subing Liu

Department of Basic Courses, Rocket Force University of Engineering, Xi'an 710025, China

* Corresponding author: Chen Chunmei (Email: 362437655@qq.com)

Abstract: In this paper, the model was established which is the test data obtained from 1214 SCL-90 mental health status test scales of a unit through factor analysis, and the evaluation results of mental health status of 1214 evaluated subjects were given by "somatization" as an example. The results of the evaluation of 10 aspects were presented, and the whole process of modeling and solving factor analysis was shown with the help of mathematical model and Matlab software. Finally, based on the analysis of the results, the comprehensive evaluation was improved by using the cluster analysis method. Not only the main causes affecting the mental health status of military personnel were found, but also the total score of each tested sample data was obtained, and a comprehensive evaluation of the mental health status of military personnel was achieved.

Keywords: Military personnel mental health, Factor analysis, Comprehensive evaluation, Cluster analysis.

1. Introduction

As the main body of troops, the mental health quality of young soldiers is an important factor affecting the combat effectiveness of troops, and improving the psychological quality of young officers and soldiers is an important task of troop construction and one of the important issues of concern to the current recruitment and troop management. Therefore, a scientific evaluation and analysis of the psychological health of young soldiers can help identify problems in time and prevent them before they occur.

For the evaluation of mental health, the SCL-90 test scale is generally used to test and analyze the test subjects. The SCL-90 examines the 90 mental health problems in 10 areas: somatization, obsessive-compulsive symptoms, interpersonal sensitivity, depressive symptoms, anxiety, hostility, fear, paranoia, psychoticism, and others. In psychology, the mental health status of the test subjects is generally evaluated by collecting data from normal people, establishing a so-called "norm", and simply comparing the mean score of each factor with the norm. However, from the perspective of mathematical modeling, if the weight of each factor in the evaluation system and the correlation between them are not considered, the evaluation of the health status lacks the persuasive power of objective evaluation. [1]

There is a require that a comprehensive evaluation model is established based on test sample data to identify various types of psychologically unhealthy individuals. The test data of self-assessment results for 1214 people in a unit are now obtained, and the format of the data is shown in Table 1. By establishing a comprehensive evaluation model, two main problems can be solved: on the one hand, it is possible to find out various types of psychologically unhealthy people in the sample data, and on the other hand, it is also possible to get which factors are the main factors on the psychological health status of military personnel among these 10 influencing factors. [2]

Table 1. Test sample data sheet of a unit

Number	Question 1	Question 2	...	Question 1214
1	2	1	...	1
2	5	2	...	1
...
1214	2	1	...	4

2. Modeling and Solving Based on Factor Analysis

There are many comprehensive evaluation methods, which can be evaluated by subjective evaluation methods: Analytic Hierarchy Process (AHP), Fuzzy comprehensive evaluation method (FUZZY), etc. or by objective evaluation methods: such as Support Vector Machine (SVM), Neural Network (NN), Discriminant Analysis, Regression Analysis, factor analysis, Cluster Analysis, Principal Component Analysis (PCA), etc.[3][4] In Table 1, the data of the collected subjects in 10 aspects such as somatization are given, which means that the objective evaluation method should be used for the evaluation on the basis that the evaluated indexes have been determined and the corresponding data have been given.

2.1. Factor Analysis Method

Factor analysis was proposed by the British psychologist Spearman in 1904, who successfully solved the statistical analysis of intelligence test scores. For a long time, educational psychologists have enriched and developed the theory and method of factor analysis and applied this method to conduct extensive research in the field of behavioral sciences. The method can represent its basic data structure by a few factors, and the weighted summation of each factor is used to obtain the combined score of each factor and construct the evaluation model of health status. Considering that the problem is the analysis and evaluation of multi-factor and multi-sample data, the factor analysis method can be used to transform the problem into a comprehensive score for each of the 10 factors and to determine the health status of each sample in each category based on the ranking results of the sample data in each aspect. [5]

2.2. Factor Analysis Model

Considering the similarity of the 10 aspects of the SCL-90 scale, including somatization, obsessive-compulsive symptoms, and interpersonal sensitivity, here we take somatization as an example for factor analysis. The somatization factors include 1, 4, 12, 27, 40, 42, 48, 49, 52, 53, 56, and 58, and we can extract a number table of 1214×12 from Table 1. The factor analysis evaluation model is built according to the following process.

Step1: Standardization of raw data

In the somatization, there are 12 variables for factor analysis, namely x_1, x_2, \dots, x_{12} , and there are 1214 evaluation objects, and the value of the j th indicator of the i th evaluation object is a_{ij} ($i=1,2,\dots,1214, j=1,2,\dots,12$). The values a_{ij} of each indicator are converted into standardized indicators \tilde{a}_{ij} ,

$$\tilde{a}_{ij} = \frac{a_{ij} - \bar{\mu}_j}{s_j}, i=1,2,\dots,1214; j=1,2,\dots,12 \quad (1)$$

Where:

$$\bar{\mu}_j = \frac{1}{1214} \sum_{i=1}^{1214} a_{ij} \quad (2)$$

$$s_j = \sqrt{\frac{1}{1214-1} \sum (a_{ij} - \bar{\mu}_j)^2} \quad (3)$$

$\bar{\mu}_j, s_j$ is the sample mean and sample standard deviation of the j th indicator.

$$\tilde{x}_j = \frac{x_j - \bar{\mu}_j}{s_j}, j=1,2,\dots,12 \quad (4)$$

\tilde{x}_j is called the standardized indicator variable.

Step2: Calculate the correlation coefficient matrix R

The correlation coefficient matrix $R = (r_{ij})_{12 \times 12}$, with

$$r_{ij} = \frac{\sum \tilde{a}_{ki} \cdot \tilde{a}_{kj}}{1214-1}, i, j=1,2,\dots,12 \quad (5)$$

where: $r_{ii} = 1, r_{ij} = r_{ji}, r_{ij}$ is the correlation coefficient of the first indicator with the first indicator.

Step3: Calculate the primary loading matrix

Calculate the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{12} \geq 0$ of the correlation coefficient matrix R, and the corresponding eigenvalue vector u_1, u_2, \dots, u_{12} , where:

$$u_j = [u_{1j}, u_{2j}, \dots, u_{12j}]^T \quad (6)$$

the primary loading matrix is

$$A_1 = [\sqrt{\lambda_1}u_1, \sqrt{\lambda_2}u_2, \dots, \sqrt{\lambda_{12}}u_{12}] \quad (7)$$

Step4: Select m primary factors

Based on the primary loading matrix, calculate the contribution of each common factor and select m principal factors. The extracted factor loadings matrix is rotated to obtain the matrix $A_2 = A_1^{(m)}T$ (where $A_1^{(m)}$ is the A_1 first m columns of and T is the orthogonal matrix).

$$\begin{aligned} \tilde{x}_1 &= \alpha_{11}F_1 + \dots + \alpha_{1m}F_m \\ &\vdots \\ \tilde{x}_{12} &= \alpha_{12,1}F_1 + \dots + \alpha_{12,m}F_m \end{aligned} \quad (8)$$

In general, the factors that play a major role can be simplified to 3-5 main factors, and we select 4 main factors.

Step5: Calculate the factor scores \hat{F}_j and evaluate them together, and use the regression method to obtain the function

$$\hat{F}_j = \beta_{j1}\tilde{x}_1 + \beta_{j2}\tilde{x}_2 + \dots + \beta_{j12}\tilde{x}_{12}, j=1,2,3,4. \quad (9)$$

Let the estimation value of the i th sample point on the j th factor score be \hat{F}_{ij} .

$$\hat{F}_{ij} = \beta_{j1}\tilde{a}_{i1} + \beta_{j2}\tilde{a}_{i2} + \dots + \beta_{j12}\tilde{a}_{i12}, i=1,2,\dots,1214, j=1,2,3,4. \quad (10)$$

Then we have

$$\begin{bmatrix} \beta_{11} & \dots & \beta_{41} \\ \vdots & \ddots & \vdots \\ \beta_{12,1} & \dots & \beta_{4,12} \end{bmatrix} = R^{-1}A_2 \quad (11)$$

And

$$\hat{F} = (\hat{F}_{ij})_{1214 \times 4} = X_0 R^{-1} A_2 \quad (12)$$

where: $X_0 = (\tilde{a}_{ij})_{1214 \times 12}$ is the standardized data matrix of the original data; R is the correlation coefficient matrix; A_2 is the loading matrix obtained in the previous step.

2.3. Model solution

From Step2 the correlation coefficient matrix R is

$$R = \begin{bmatrix} 1.00 & 0.44 & 0.29 & 0.32 & 0.32 & 0.36 & 0.37 & 0.39 & 0.32 & 0.36 & 0.30 & 0.37 \\ 0.44 & 1.00 & 0.39 & 0.40 & 0.46 & 0.31 & 0.44 & 0.49 & 0.48 & 0.39 & 0.41 & 0.41 \\ 0.29 & 0.39 & 1.00 & 0.38 & 0.46 & 0.40 & 0.62 & 0.47 & 0.49 & 0.48 & 0.46 & 0.51 \\ 0.32 & 0.40 & 0.38 & 1.00 & 0.54 & 0.52 & 0.55 & 0.55 & 0.54 & 0.53 & 0.48 & 0.46 \\ 0.32 & 0.46 & 0.46 & 0.54 & 1.00 & 0.38 & 0.50 & 0.55 & 0.51 & 0.49 & 0.48 & 0.46 \\ 0.36 & 0.31 & 0.40 & 0.52 & 0.38 & 1.00 & 0.54 & 0.56 & 0.53 & 0.48 & 0.56 & 0.54 \\ 0.37 & 0.44 & 0.62 & 0.55 & 0.50 & 0.54 & 1.00 & 0.62 & 0.63 & 0.61 & 0.52 & 0.59 \\ 0.39 & 0.49 & 0.47 & 0.55 & 0.55 & 0.56 & 0.62 & 1.00 & 0.64 & 0.58 & 0.60 & 0.59 \\ 0.32 & 0.48 & 0.49 & 0.54 & 0.51 & 0.53 & 0.63 & 0.64 & 1.00 & 0.60 & 0.61 & 0.58 \\ 0.36 & 0.39 & 0.48 & 0.53 & 0.49 & 0.48 & 0.61 & 0.58 & 0.60 & 1.00 & 0.56 & 0.54 \\ 0.30 & 0.41 & 0.46 & 0.48 & 0.48 & 0.56 & 0.52 & 0.60 & 0.61 & 0.56 & 1.00 & 0.57 \\ 0.37 & 0.41 & 0.51 & 0.46 & 0.46 & 0.54 & 0.59 & 0.59 & 0.58 & 0.54 & 0.57 & 1.00 \end{bmatrix}$$

From Step3, the primary loadings matrix A_1 is

$$A_1 = \begin{bmatrix} 0.53 & 0.68 & 0.36 & 0.16 & 0.19 & -0.09 & 0.16 & 0.04 & -0.04 & -0.16 & 0.02 & 0.01 \\ 0.63 & 0.53 & -0.24 & -0.11 & -0.34 & 0.08 & -0.28 & 0.05 & 0.12 & 0.18 & 0.00 & -0.02 \\ 0.68 & -0.07 & -0.37 & 0.49 & 0.18 & 0.17 & -0.02 & 0.22 & 0.03 & -0.04 & -0.06 & 0.20 \\ 0.72 & -0.07 & 0.04 & -0.44 & 0.34 & 0.08 & -0.22 & 0.00 & 0.24 & -0.18 & -0.14 & 0.07 \\ 0.70 & 0.09 & -0.37 & -0.32 & 0.16 & 0.16 & 0.39 & -0.03 & -0.11 & 0.08 & 0.18 & -0.07 \\ 0.71 & -0.18 & 0.48 & 0.00 & 0.07 & 0.29 & -0.10 & 0.15 & -0.09 & 0.27 & 0.16 & 0.03 \\ 0.81 & -0.09 & -0.10 & 0.21 & 0.19 & -0.08 & -0.22 & -0.10 & -0.13 & -0.01 & -0.02 & -0.40 \\ 0.81 & -0.03 & 0.04 & -0.10 & -0.13 & 0.01 & 0.03 & -0.16 & -0.36 & 0.04 & -0.37 & 0.14 \\ 0.80 & -0.14 & -0.03 & -0.05 & -0.21 & -0.15 & -0.16 & -0.05 & -0.17 & -0.28 & 0.33 & 0.14 \\ 0.76 & -0.12 & -0.01 & 0.01 & 0.09 & -0.54 & 0.07 & 0.07 & 0.12 & 0.26 & 0.01 & 0.09 \\ 0.76 & -0.20 & 0.12 & -0.04 & -0.34 & 0.01 & 0.23 & 0.32 & 0.13 & -0.17 & -0.12 & -0.18 \\ 0.77 & -0.10 & 0.10 & 0.23 & -0.13 & 0.13 & 0.15 & -0.44 & 0.30 & 0.01 & 0.03 & 0.03 \end{bmatrix}$$

The contribution rates of the four principal factors found are shown in Table 2.

Table 2. Contribution rates of the principal factors

Factors	1	2	3	4
The contribution rates	26.95	11.03	17.52	16.80
Cumulative contribution rate	26.95	37.98	55.50	72.30

In Step5, the contribution from

$$\begin{bmatrix} \beta_{11} & \cdots & \beta_{41} \\ \vdots & \ddots & \vdots \\ \beta_{12,1} & \cdots & \beta_{4,12} \end{bmatrix} = R^{-1}A_2 \quad (13)$$

obtained, and

The resulting score function for each factor is

$$R^{-1}A_2 = \begin{bmatrix} -0.0134 & 0.9122 & 0.2211 & -0.1326 \\ -0.3702 & 0.3991 & -0.4429 & 0.0174 \\ -0.2836 & -0.0691 & 0.1646 & 0.8235 \\ 0.2138 & -0.1415 & -0.4337 & -0.4209 \\ -0.2158 & -0.1547 & -0.6515 & -0.0512 \\ 0.5626 & 0.0852 & 0.3413 & -0.2423 \\ 0.0133 & -0.0368 & 0.0903 & 0.3604 \\ 0.1536 & -0.0038 & -0.1159 & -0.0630 \\ 0.1316 & -0.1287 & -0.1068 & 0.0488 \\ 0.1352 & -0.0891 & -0.0491 & 0.0660 \\ 0.2849 & -0.1150 & 0.0358 & -0.0378 \\ 0.1733 & 0.0581 & 0.2620 & 0.2460 \end{bmatrix} \quad (14)$$

The resulting score function for each factor is

$$\begin{aligned} F_{11} &= -0.0134\tilde{x}_1 - 0.3702\tilde{x}_2 - 0.2836\tilde{x}_3 + 0.2138\tilde{x}_4 \\ &\quad - 0.2158\tilde{x}_5 + 0.5626\tilde{x}_6 + 0.0133\tilde{x}_7 + 0.1536\tilde{x}_8 \\ &\quad + 0.1316\tilde{x}_9 + 0.1352\tilde{x}_{10} + 0.2849\tilde{x}_{11} + 0.1733\tilde{x}_{12} \\ F_{12} &= 0.9122\tilde{x}_1 + 0.3991\tilde{x}_2 - 0.0691\tilde{x}_3 - 0.1415\tilde{x}_4 \\ &\quad - 0.1547\tilde{x}_5 + 0.0852\tilde{x}_6 - 0.0368\tilde{x}_7 - 0.0038\tilde{x}_8 \\ &\quad - 0.1287\tilde{x}_9 - 0.0891\tilde{x}_{10} - 0.1150\tilde{x}_{11} + 0.0581\tilde{x}_{12} \\ F_{13} &= 0.2211\tilde{x}_1 - 0.4429\tilde{x}_2 + 0.1646\tilde{x}_3 - 0.4337\tilde{x}_4 \\ &\quad - 0.615\tilde{x}_5 + 0.3413\tilde{x}_6 + 0.0903\tilde{x}_7 - 0.1159\tilde{x}_8 \\ &\quad - 0.1068\tilde{x}_9 - 0.0491\tilde{x}_{10} + 0.0358\tilde{x}_{11} + 0.2620\tilde{x}_{12} \\ F_{14} &= -0.1326\tilde{x}_1 + 0.0174\tilde{x}_2 + 0.8235\tilde{x}_3 - 0.4209\tilde{x}_4 \\ &\quad - 0.0512\tilde{x}_5 - 0.2423\tilde{x}_6 + 0.3604\tilde{x}_7 - 0.0630\tilde{x}_8 \\ &\quad + 0.0488\tilde{x}_9 + 0.0660\tilde{x}_{10} - 0.0378\tilde{x}_{11} + 0.2460\tilde{x}_{12} \end{aligned}$$

Using the composite factor score formula:

$$F_1 = \frac{26.95F_{11} + 11.03F_{12} + 17.52F_{13} + 16.80F_{14}}{72.30} \quad (15)$$

Matlab was used to calculate the composite scores of the somatization factors of the 1214 military personnel psychological conditions, which are shown in Table 3.

Table 3. Composite scores of somatization factors

Number	1	2	3	4
F_i	0.319	0.073	0.724	-0.244
Number	1212	1213	1214
F_i	0.145	-0.075	-0.104

The histogram of the frequency distribution of the scores is drawn in Figure 1, through which it can be seen that the combined scores of the 1214 test samples have a roughly normal distribution trend.

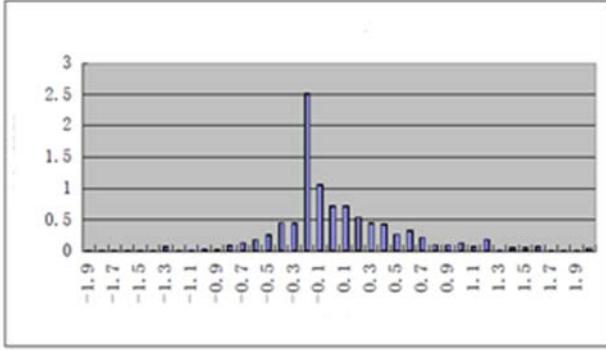


Figure 1. Histogram of somatization frequency distribution

The other 9 types of mental health problems evaluation score procedures only need to change the somatization data in the above procedures to the respective data files corresponding to the factors to be evaluated: obsessive-compulsive symptoms, interpersonal sensitivity, depression, anxiety, hostility, terror, paranoia, psychoticism, and others.

3. Results

Through the factor analysis method, the processing and analysis difficulties brought by the large amount of data are solved, and the subjective arbitrariness brought by setting the weights of each index in the hierarchical analysis is avoided, and the evaluation method of comparing the mean with the normative comparison is more convincing and practical. Because the factor analysis method analyzes the 1214 test data from 10 aspects separately, each aspect is able to obtain a ranking result, which enables the evaluator to visually evaluate the mental health status of the 1214 evaluated subjects from multiple aspects.

Although we have completed the evaluation of all aspects of mental health separately by factor analysis, we also face a problem that the evaluation of the mental health of the 1214 subjects is not easy. However, we also face a problem that we need to conduct a comprehensive evaluation of the psychological status of the 1214 subjects. The purpose of the comprehensive evaluation is to find the dominant factors among the 10 factors that affect the psychological health of military personnel, and to analyze the main reasons that affect the psychological health of military personnel. Due to the high correlation between variables, multiple variables lead to a large amount of data, which brings a lot of inconvenience to the analysis of the system. Therefore, we used the method of cluster analysis to aggregate the factors affecting the mental health of military personnel into fewer categories according to the similarity of the variables, and then identify the main causes affecting the system.

4. Improvement of the Model

The main process of R-type cluster analysis (clustering of variable indicators) is as follows.

Step1: Standardization process. (Same as the standardization method of factor analysis)

Step2: Variable indicator similarity measure

In the cluster analysis of variables, the first step is to determine the similarity measure of variables, and we use the correlation coefficient.

Note the value of the variable indicator. Then the sample correlation coefficients of the two variables and can be used as their similarity measures, i.e.

$$r_{kl} = \frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)(x_{il} - \bar{x}_l)}{\left[\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2 \sum_{i=1}^n (x_{il} - \bar{x}_l)^2 \right]^{\frac{1}{2}}} \quad (16)$$

The correlation coefficient matrix R can be calculated,

$$R = (r_{kl})_{10 \times 10}.$$

Step3: Variable clustering method

The variable clustering method adopts the same idea and process as the systematic clustering method, and here we use the longest distance method for clustering, defining the distance between the two variables as

$$R(G_1, G_2) = \max_{\substack{x_k \in G_1 \\ x_l \in G_2}} \{d_{kl}\} \quad (17)$$

where: $d_{kl} = 1 - |r_{kl}|$, $R(G_1, G_2)$ is related to the similarity measure between the two variables with the least similarity in the two classes.

Step4: Use Matlab to draw the clustering diagram and get the values γ_i of each factor scale.

Step5: Analyze the clustering diagram and calculate the weight A matrix of each influence factor. where the elements a_i of A are

$$a_i = \frac{\gamma_i}{\sum_{i=1}^n \gamma_i}, i = 1, 2, \dots, n \quad (18)$$

Step6: Calculate the total score M of each tester's mental health status.

$$M_{1214 \times 1} = F_{1214 \times 10} \times A_{10 \times 1} \quad (19)$$

where $F_{1214 \times 10}$ is the matrix of the combined scores of each factor of the testers.

The clustering diagram of the causes affecting the mental health of military personnel was drawn using Matlab, as shown in Figure 2.

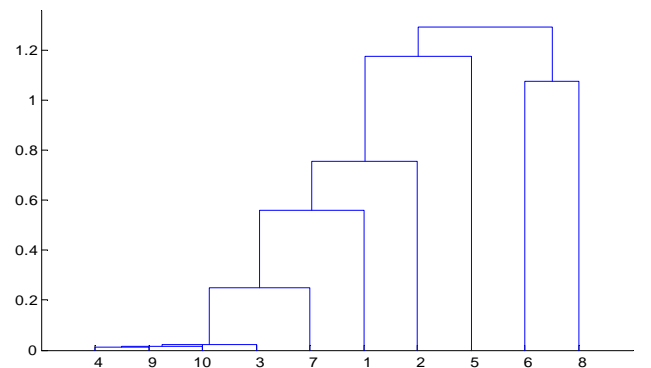


Figure 2. Clustering diagram of military mental health

The values γ_i of the scales obtained from the clustering diagram are shown in Table 4.

Table 4. The values of impact factors of military mental health

factors	1	2	3	4	5
values	0.582	0.846	0.235	0.157	0.680
factors	6	7	8	9	10
values	0.679	0.546	0.785	0.157	0.194

Calculation of the weights of each influence factor is shown in Table 5.

Table 5. Weights of impact factors of military mental health

factors	1	2	3	4	5
weights	0.123	0.179	0.050	0.033	0.144
factors	6	7	8	9	10
weights	0.144	0.116	0.166	0.033	0.041

According to the results of the cluster analysis in Figure 3, and the scale values and weights of the 10 factors in Tables 4 and 5, we can know that factors 2, 8, 5 and 6 have the greatest influence on the mental health of military personnel, that is, the four factors of obsessive-compulsive symptoms, paranoia, hostility and anxiety are the main causes of the mental health status of military personnel.

It is reasonable to further analyze these 4 factors as the main reasons affecting the mental health of military personnel from their identity characteristics and occupational characteristics. Compulsive symptoms are due to the army's demand for absolute obedience, so for soldiers, whether they are willing to do a certain task arranged by their superiors or not, they must do what they are told, and they will also be required to do as well as possible and strive for excellence. Hostility is due to the fact that the troops are, after all, to fight the war, the competition is relatively more intense. Always war as their future hypothetical life, so it is normal to produce hostility in this environment. Paranoia and anxiety become the main reason because the soldiers are in a relatively closed environment, and there is a specific day-to-day life system, if problems are not corrected and guided in time, it is easy to have anxiety in the mind and cause paranoid views.

Finally, the total score of mental health status of each tester was calculated as shown in Table 6.

Table 6. Comprehensive evaluation score of military mental health

Number	1	2	3	4	5
Scores	0.230	0.106	0.309	0.363	0.254
Number	1211	1212	1213	1214
Scores	0.709	0.139	0.332	0.298

Using the cluster analysis method, we not only found the main causes affecting the mental health status of military personnel, but also obtained the total score of each tested sample data, and achieved a comprehensive evaluation of the mental health status of military personnel.

5. Conclusion

In this paper, factor analysis is used to obtain the comprehensive evaluation results of 10 factors affecting the mental health status of military personnel, and the evaluation of all aspects of mental health separately is completed. on this basis, cluster analysis is adopted to find the main causes affecting the mental health status of military personnel, and the total score of each tested sample data is obtained. After comprehensive evaluation, the same test sample can not only know its situation in a certain factor, but also get its comprehensive situation in all factors, which effectively ensures the objective, comprehensive and valid evaluation results.

References

- [1] Yang Zhongliang, Wang Changshan, Liu Boshi, et al. Mental Health Examination of 3017 Recruits and Establishment of a Comprehensive Assessment Model for Recruits Mental Health[J].China Journal of Health Psychology, 2015, 23(10): 1471-1474.
- [2] Feng Zhengzhi, Zu Xia. Mental health assessment for military men:theories and model[J]. Journal of the Third Military Medical University, 2015, 37(22):2207-2212.
- [3] Wang Jian, Zhao Guosheng. Matlab Mathematical Modeling and Simulation[M]. Beijing: Tsinghua University Press, 2016, pp. 479-489.
- [4] He Xiaoqun. Multivariate Statistical Analysis[M]. Beijing: Renmin University Press, 2015, pp. 142-170.
- [5] Si Shoukui, Sun Xijing. Mathematical Modeling Algorithms and Applications[M]. Beijing: National Defense Industry Press, 2019, pp. 234-255.