Research on Mental Health Status and Influence Factors of College Students Based on Multivariate Analysis and Bayesian Discrimination

Lixin Deng
Shanghai University of Electric Power
Shanghai, China
d_eng_lixin@126.com

Keywords: multivariate analysis; bayesian discrimination; multivariate variance analysis; linear best solution; independent variables; highly correlated variables

Abstract. In recent years, with the accelerating pace of society, mental health problems of college students concerned by the communities widely. Due to the psychology of college students is not mature, campus accidents caused by psychological problems are increased every year. So, the research and analysis on the mental health of college students are very necessary. Firstly, this paper establishes college mental health variables analysis model using multivariate data analysis and puts forward the math regression model of independent variable and highly correlated variables. Then, it finds out the main factors that affect the mental health of college students by multivariate Bayesian discrimination, and puts forward the best solution of linear multivariate analysis of variance. Finally, it substitutes all the statistical data into the SPSS software and does the mathematical multivariate analysis, and gives the detailed calculation steps and results. It analyzes the results of curve which proved the reliability and stability of the analysis of psychological conditions and provides a reliable theoretical foundation for the college students' mental status and influencing factors.

Introduction

Psychological development of college students are influenced by many factors including school, family, and society. With the accelerating pace of society, diversification trend of development of society is increasingly evidently. As the new generation of college students, psychological development becomes more and more complicated under the influence of the community[1]. Due to the improvement of economic conditions of college students' families, students’ psychologies have great influenced by consumption concept, network environment, social atmosphere and moral values[2,3]. This paper analyzes the factors that affect the mental health of college students through 1000 questionnaires. The first part of this paper gives multivariate Bayesian analysis of mental health model of college students and the discriminative statistical functions of the mental health of college students. The second part of the paper processes the results of the analysis by the SPSS software, and optimizes the results through statistical analysis and gets linear deviation fitting curve. At last, it concludes the main factors of the mental health of college students and provides a reference for the development of the mental health of college students[4-7].

Multivariate Bayesian discrimination analysis model of mental health of college students

The multivariate Bayesian discrimination method not only takes the result of misjudgment analysis into account but also takes priori frequency before the calculation into account[8-10]. Multivariate Bayesian discrimination method combines the probability distribution with priori frequency and gets the posterior frequency. Multivariate Bayesian discrimination analysis of mental health of college students can get the influencing factors of mental health of college students through posterior frequency. Assuming that the probability of the mental health of college students is \( P(D) \), the probability of mental unhealthy of college students is \( P(F) \) and the formula is as follows[11]:

\[
P(F_i / D) = \frac{P(D / F_i) P(F_i)}{\sum P(D / F_i) P(F_i)}
\]

\[ (1) \]
Assuming that the overall survey is Hj (j=1, 2... n), Probability density function of Hj is \( m_i(a) \)
Assuming that the probability of Hj is \( k_j \). a is the event of mental health or mental unhealthy. When a occurs, we can find the category which it belongs. Bayesian multivariate discrimination of posterior probability formula can be expressed as follows[12]:

\[
P(H_j \mid a) = \frac{m_i(a)k_j}{\sum_{i=1}^{n}m_i(a)}
\]

(2)

Multiple discrimination rules can be expressed as follows[13]:

\[
P(H_j \mid a) = \frac{k_j}{\sum_{i=1}^{n}k_i} = \max_{i \in \{1, 2, ..., n\}} \frac{k_j}{\sum_{i=1}^{n}k_i}
\]

(3)

Under the assumption of normal distribution, we can assume \( m_i(a) \) is density function of normal distribution;

\[
k_i(a) = \max_{i \in \{1, 2, ..., n\}} k_i(a)
\]

(4)

If it can meet (4), we can know it is health. If

\[
m_i(a) = \frac{1}{(2\pi)^{d/2}\left|\Sigma_j\right|^{1/2}} \exp\left[-\frac{1}{2}(a - u^{(j)})\Sigma_j^{-1}(a - u^{(j)})\right]
\]

(5)

\[
K_j(a) = k_j \frac{1}{(2\pi)^{d/2}\left|\Sigma_j\right|^{1/2}} \exp\left[-\frac{1}{2}(a - u^{(j)})\Sigma_j^{-1}(a - u^{(j)})\right]
\]

(6)

So,

Through (5) , (6), We can calculate the discrimination function[14]:

\[
y_j(a) = \ln(k_i(a))
\]

(7)

When \( Z_j(x) = \max_{i \in \{1, 2, ..., n\}} Z_j(x) \), it is healthy.

If covariance matrix is \( \Sigma_i = \cdots = \Sigma_k = \Sigma \), the discrimination function can expressed as:

\[
y_j(a) = \ln k_j + \frac{1}{2}(a - u^{(j)})\Sigma^{-1}(a - u^{(j)})
\]

(8)

\[
= -\frac{1}{2}[-2\ln k_j + (a - u^{(j)})\Sigma^{-1}(a - u^{(j)})]
\]

If

\[
V_j(a) = -2\ln k_j + (a - u^{(j)})\Sigma^{-1}(a - u^{(j)})
\]

\[
Z_j(a) = -2\ln k_j - 2u^{(j)}\Sigma^{-1}a + u^{(j)}\Sigma^{-1}u^{(j)}
\]

(9)

When \( Z_j(a) = \min_{i \in \{1, 2, ..., n\}} Z_j(a) \), it is healthy.

**SPSS software multivariate Bayesian discrimination analysis**

This paper surveys mental health problems of college students using questionnaires. The contents mainly are situations of mental health of college students influenced by family education, school education, social trends, network environment, and consumption concept, the scope of the investigation is shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE I. SITUATION OF VALID QUESTIONNAIRES</th>
<th>TABLE II. SURVEY TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Range</strong></td>
<td><strong>Network interactive content</strong></td>
</tr>
<tr>
<td>Bachelor</td>
<td>Times</td>
</tr>
<tr>
<td>500</td>
<td>300</td>
</tr>
<tr>
<td>Professional school</td>
<td>300</td>
</tr>
<tr>
<td>Higher vocational school</td>
<td>100</td>
</tr>
<tr>
<td>School of art</td>
<td>100</td>
</tr>
</tbody>
</table>

In order to increase the universality and validity of questionnaires, this paper processes the questionnaire, and eliminates invalid questionnaires, and increases the relevance of the survey[15]. From Table 1, we can see that valid questionnaires of bachelors and professional schools occupy a large proportion. The proportion of bachelors and professional schools in all colleges and
universities is large. The proportion of vocational high schools and art schools in all colleges and universities is small.

Table 2 is the times and frequencies of different content surveys. It can be seen from Table 2 that the frequency of family education is the highest because Family education is the main influencing factors of students' psychological development, and family education plays a dominant role in the child's psychological growth process. Just a part of students whose economic conditions is good have consumption concept. A lot of college students whose family conditions is better like extravagance, waste and comparison in the consumer psychology which influence the development of mental health but not the main factors.

Symptom Checklist SCL-90 is one of the most famous test scales of mental health in the world. It is widely used. The paper analyzes statistical results using SPSS, and gets the statistical table of mental health of college students which is shown in Table 3.

The process of SPSS software statistical analysis needs to analyze MDA mode. The covariance is the main factor influencing the accuracy of discrimination. The actual operation process of multivariate analysis of variance is shown as follows[16,17]:

1. We can open the analyze module, and select the General Lineal model and then select Repeated measures;
2. We can write fangcha as subject in within-subject factor name;
3. We can write 3 in the title box of the number of levels; (4) We can select ADD button;
4. We can select DEFINE button;
5. We can select fangcha1-4 in title box of within-subject variables;
6. We can select the subject, anxiety and tension in between subjects facto dialog box;
7. We can select MODEL button; (9) We can select Custom button;
8. We can write fangcha in Within-subject Model;
9. We can choose anxiety and tension in Between subjects Model;
10. We can select CONTINUE button; (13) We can select OK.

We can get variance curve through the analysis of SPSS variance which is shown as follows.

From the variance fitting curve of Figure 1, we can see that the result of multiple measurements in the multiple discriminant is a straight line which proves the error presented linear best. The share ratio of mental health influencing factors can be obtained after optimization of linear best which is shown in Table 4.

<table>
<thead>
<tr>
<th>Psychological symptoms</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obsessive-compulsive disorder status</td>
<td>2.85</td>
<td>0.26</td>
</tr>
<tr>
<td>Depression status</td>
<td>2.12</td>
<td>0.32</td>
</tr>
<tr>
<td>Situation of anxiety disorders</td>
<td>3.56</td>
<td>0.25</td>
</tr>
<tr>
<td>Phobia condition</td>
<td>0.38</td>
<td>0.02</td>
</tr>
<tr>
<td>Paranoia conditions</td>
<td>1.58</td>
<td>0.15</td>
</tr>
</tbody>
</table>

TABLE III. THE SITUATION OF MENTAL HEALTH

<table>
<thead>
<tr>
<th>Network interactive content</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Education</td>
<td>28%</td>
</tr>
<tr>
<td>School education</td>
<td>22%</td>
</tr>
<tr>
<td>Social atmosphere</td>
<td>20%</td>
</tr>
<tr>
<td>Consumption concept</td>
<td>12%</td>
</tr>
<tr>
<td>Network environment</td>
<td>18%</td>
</tr>
</tbody>
</table>

TABLE IV. THE PROPORTION OF MENTAL HEALTH IMPACT FACTORS

Figure 1. The variance fitting curves

Figure 2. The proportion of mental health impact factors

It can be seen from Table 4 and Figure 2, in the factors which influence college mental health, the proportion of family education is the largest, school education is the second largest, and network environment is the third largest. Family education and school education should be conducted mainly in the psychological health education of college students. We should guide students to develop a positive mental. Network environment is one of the main factors affect the mental health of college students.
students. So, we should concerned about the development of students' psychological factors in network environment, and create healthy and progressive environment.

**Conclusion**

This paper researched the mental health situations and the main influencing factors of college students using multivariate analysis method. In the first part of the text, this paper gave Bayesian discrimination model of college students' mental health survey. Mental health discrimination probability function is given by the model. It distinguished the degree of mental health by covariance. The second part analyzed five factors that affect the mental health of college students in the form of questionnaires. It optimized the questionnaires contents accordance with the scope of the investigation and survey frequency. Finally, it used SPSS software to deal with the statistical results of the 1000 survey questionnaire and gave the SPSS detailed steps of error optimization. After the error fitting curve process, we got the factors and the proportion that influence the development of mental health of college students. From the results, we found that family education, school education and social values and the network environment are the main factors impact the development of mental health of college students in all factors. School, families and the community should work together to create a positive growth environment for the development of mental health of college students which ensure the development of mental health of college students.

**References**


