Analysis of the relationship between the sleep quality of each index based on the stepwise regression model

Junjie Xiao [a], Xue Lu [b] and Haixin Zhou [c]

School of North China University of Science and Technology, Tangshan 063200, China;

[a]1114163822@qq.com, [b]1417176692@qq.com, [c]962818055@qq.com

Abstract

The quality of sleep affects the efficiency of work and the quality of the body, and there are many factors that affect the quality of sleep.In this paper, the relationship between sleep quality and quality of sleep was analyzed by the quality of sleep given in the middle Annex 2 given by Asia and Pacific Mathematical Contest in Modeling and the various indexes.First, In order to find out the relationship between indicators and sleep quality, through the correlation analysis of each index, it is concluded that there is a strong relationship between each index and sleep quality.Introduce the stepwise regression model .Stepwise regression analysis was carried out with 6 indicators as variables, exclude the two indexes (reliability and mental state) which had little correlation with sleep quality, and the regression equation was obtained: $y = 1.343 + 0.648x_1 + 0.100x_2 + 0.658x_5 - 0.106x_6$ .It can be seen that age, sex and tension are positively correlated with sleep quality, and personality is negatively related to sleep quality.

Keywords

Sleep quality, Poisson correlation, stepwise regression.

1. Introduction

Since 2001, the World Association of Sleep Medicine has designated March 21 of each year as World Sleep Day to remind people of the importance and quality of sleep. The mental state of a day depends on the quality of sleep last night, high sleep quality will naturally make people energetic, with the increasing pressure of learning life, insomnia is also rising, long-term insomnia can make people feel tired, energy is not Concentration, resulting in low efficiency of work and study, severe insomnia and even cause autonomic dysfunction, leading to various problems such as system imbalance in the body. In this paper, the relationship between the quality of sleep and the quality of sleep was analyzed by the quality of sleep given in Annex 2 inAsia and Pacific Mathematical Contest in Modelingand the various indexes.

2. The establishment and solution of the model

2.1 Correlation analysis

Analyze the relationship between sleep quality and each index, equivalent to study the correlation between ies and sleep quality of human body, Analysis of indicators related to each other, to measure the correlation bndexetween each index and the quality of human sleep. Statistical software was used to get the correlation coefficient table between each index and human sleep quality. As shown in Table 1:

Table 1 Index correlation coefficient table

Pearson correlation	Age	Sex	Sleep quality	Reliability	Psychoticism	Nervousness	Character
Age	1	0.048	0.130	0.384	0.210	-0.250	0.080
Sex	0.048	1	0.058	0.048	0.351	-0.043	0.057
Sleep quality	0.130	0.058	1	0.017	0.078	0.083	-0.031
Reliability	0.384	0.048	0.017	1	-0.182	-0.355	-0.002
Psychoticism	0.210	0.351	0.078	-0.182	1	0.039	0.019
Nervousness	-0.250	-0.043	0.083	-0.355	0.039	1	-0.185
Character	0.80	0.057	-0.031	-0.002	0.019	-0.185	1

According to the correlation coefficient matrix in the graph, it can be seen preliminarily, age, gender, reliability, mental status and stress level were positively correlated with sleep quality. The correlation from strong to weak is: age, tension, mental state, gender, reliability, personality and human sleep quality is negatively related. Overall, there is a linear relationship between indicators and sleep quality. In order to further explore the quantitative relationship between each index and sleep quality and its significance, the stepwise regression model was introduced to select and eliminate the independent variable of multiple linear regression equation, and then the quantitative relationship between human sleep quality and indicators and each index was obtained.

2.2 Fundamentals - Stepwise regression

The basic idea of stepwise regression[1] is to introduce the variables one by one into the model. Each introduction of an explanatory variable should be tested, and the explanatory variables that have been selected must be tested one by one. When the originally introduced explanatory variable becomes invalid due to the introduction of later explanatory variables When it is significant, delete it. To ensure that the regression equation contains only significant variables before each new variable is introduced. The above steps are repeated until no significant explanatory variable is entered into the regression equation or no insignificant explanatory variable is removed from the regression equation.

There are m independent variables, using m independent variables to model the full model, that is:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{\mathrm{m}} x_m + \varepsilon \tag{1}$$

Remove the independent variables from the m variables and use the m-1 independent variables fitting model as the subtraction model, that is:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} \cdots + \beta_m x_m + \varepsilon \tag{2}$$

The whole model of the complex judgment coefficient is R^2 , minus the model of complex judgment coefficient of $\mathrm{R}_{\mathrm{j}}^2$ .definition:

$\Delta \mathrm{R}_{\mathrm{j}}^2 = \mathrm{R}^2 - \mathrm{R}_{\mathrm{j}}^2$

Since there is more than one argument x_j in the whole model,Therefore, if $\Delta \mathrm{R}_{\mathrm{j}}^2$ is almost zero, it indicates that the increase of x_j has no significant improvement on the explanatory power y of the pair; and if $\Delta \mathrm{R}_{\mathrm{j}}^2$ is significantly non-zero, x_j can provide significant explanatory information for the regression model.

Give statistical assumptions H_0: $\Delta \mathrm{R}_{\mathrm{j}}^2 = 0$, H_1: $\Delta \mathrm{R}_{\mathrm{j}}^2 \neq 0$

The statistical test quantity is: $\mathrm{F}_{\mathrm{j}} = \dfrac{\mathrm{Q}_{\mathrm{j}} - \mathrm{Q}}{\mathrm{Q}/(\mathrm{n-m-1})}$

Among them, Q_{j} is the sum of squares of the residuals subtracted from the model and Q is the square sum of the residuals of the full model.

According to the inspection level check F distribution table, get the critical value of the rejection domain, the decision criteria are:

When $F_j > F_\alpha$, denied H_0, the notation ΔR_j^2 is not significant, indicating that the introduction of variables $x_1, \cdots, x_{j-1}, x_{j+1}, \cdots, x_m$ will significantly improve the ability to interpret the pair.

(2)When $F_j \le F_\alpha$, Acceptable H_0, ΔR_j^2 is significantly zero, so removing x_j from the full model shows no noticeable diminishing changes in the ability to interpret y.

The obtained multiple linear regression equation is significant, the overall parameters need to be significant test,

Test statistics are:

$$t_0 = \frac{\hat{\beta}_0}{S\left(\hat{\beta}_0\right)}$$

β_0 confidence interval of $1-\alpha$ confidence intervals:

$$\hat{\beta}_0 - t_{\alpha/2}(n-2)S\left(\hat{\beta}_0\right) \le \beta_0 \le \hat{\beta}_0 + t_{\alpha/2}(n-2)S\left(\hat{\beta}_0\right)$$

The multivariate regression equation of the parameters tested, if obtained within the confidence interval, indicating that the equation is established without problems.

2.3 Stepwise regression model establishment

Quantification of sleep quality indicators[2]: In sleep indicators, age and gender are qualitative data, quantify qualitative data with quantitative data of other indicators to make full use of these information to more fully explore the relationship between human sleep quality and various indicators. Qualitative data in the original data to quantify and unified processing, data processing is as follows:

$$age = \begin{cases} 0, & < 40 \\ 0.25, & [40,50) \\ 0.5, & [50,60) \\ 0.75, & [60,70) \\ 1, & \ge 70 \end{cases}$$

$$sex = \begin{cases} 0, & male \\ 1, & female \end{cases}$$

To human sleep quality as an explanatory variable, the body may affect the quality of sleep indicators as explanatory variables, the establishment of multiple linear regression equation, according to Figure 1 available equation is as follows:

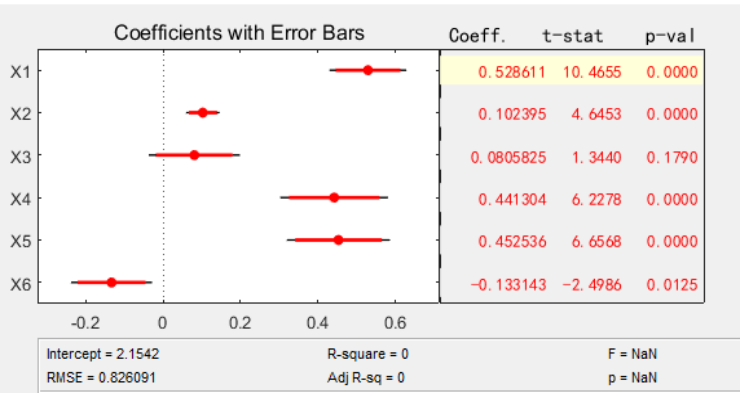

Fig. 1 Multiple linear regression equation coefficient chart

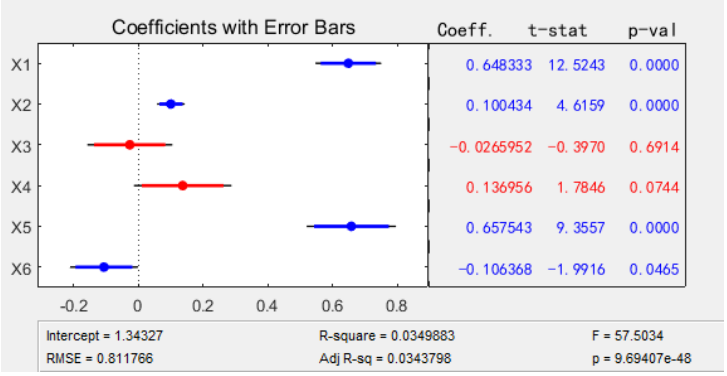

Fig. 2 Step-by-step regression equation coefficient chart

After three regression, from figure 1 to figure 2, It can be seen in figure 2 that the value P of x_1, x_2, x_5, x_6 all less than 0.05, the value P of x_3, x_4 all greater than 0.05, the regression coefficients of x_3 and x_4 were not significant, in order to further explore the impact of various indicators on human sleep quality and correlation, stepwise regression method was used to determine the quantitative relationship between indicators and sleep quality, and the significant weak indicators x_3 and x_4 were eliminated. According to figure 2, the equation was as follows:

$$y = 1.3433 + 0.6483x_1 + 0.1004x_2 + 0.6575x_5 - 0.1064x_6 \quad (3)$$

3. Conclusion

According to the regression equation and stepwise regression process, the reliability and mental state of these two indicators and sleep quality is relatively small, the impact on the quality of sleep is very low, it will be removed. Effect of age, gender, anxiety, personality and quality of sleep has a certain effect, which has significant positive effect on sleep quality with age, degree of tension, gender and personality have influence on the quality of sleep, but the effect is not significant, and the character has a negative impact on sleep quality.

References

[1] He Xiaoqun. Beijing: Higher Education Press, Practical Regression Analysis, 2011.6.30.
[2] Shen Huayu. Application of neural network in medical diagnosis [D], Tianjin University of Technology, 2007.
[3] Shao Jun. Design of Traffic Management System Based on Stepwise Regression Prediction Model [F]. Fudan University, 2013.