ISSN : 0974 - 7435
Volume 10 Issue 9
2014
BioTechnology
An Indian Journal
FULL PAPER
BTAIJ, 10(9), 2014 [3765-3771]

Variation phenomenon of group behavior in bioinformatics

Qingmin Kong, Xiuqing Liang
School of Business, Guangxi University, Nanning, Guangxi, 530004, (P.R.CHINA)

ABSTRACT

Group behavior is one of the most important and common tasks in bioinformatics, however, it is still not clear that whether variation phenomenon exist in group behavior. The aim of this study is to find out the variation phenomenon in bioinformatics. In this study, factor analysis algorithm and variance analysis algorithm is used to calculate variation phenomenon of group behavior in bioinformatics. Cultural behavior is used as sample for simulations. After simulations, the significance of variation phenomenon has been found. This study can be directions for future bioinformatics research and development.

KEYWORDS

Variation phenomenon; Bioinformatics; Factor analysis algorithm; Variance analysis algorithm; Group behavior; Culture.

INTRODUCTION

Group behavior is a hot issue in the study of bioinformatics. Some scholars found that group spirit and thought can form and predict individual behavior (Song, 2008)[1]. Group behavior generates individual behavior, especially in cultural behavior. There are several cultural behavior have been found (Hofstede, 2001; Anderson, 2002; Smith, 2004)[2-4]. However, it is still not clear that what variation phenomenon should have in bioinformatics.

The aim of this study is to find out the variation phenomenon in bioinformatics. Cultural behavior is used as sample for simulations to test the variation phenomenon in bioinformatics.

MATERIALS AND METHODS

This study analyzes the variation phenomenon in bioinformatics, in which cultural behavior is used as sample.

The research procedures were shown in Figure 1.

The core of factor analysis is to show most information of original variables through a few independent factors (Xue, 2006; Harman, 1967)[5,6]. We suppose there are several original variables x1, x2, x3, …, xp, each variable is 0.000 mean and 1.000 standard deviation. The original variable can be expressed as a linear combination by k (k<p) factors, such as:

$$\begin{cases} x_1 = \alpha_{11} f_1 + \alpha_{12} f_2 + \alpha_{13} f_3 + \cdots + \alpha_{1k} f_k + \varepsilon_1 \\ x_2 = \alpha_{21} f_1 + \alpha_{22} f_2 + \alpha_{23} f_3 + \cdots + \alpha_{2k} f_k + \varepsilon_2 \\ x_3 = \alpha_{31} f_1 + \alpha_{32} f_2 + \alpha_{33} f_3 + \cdots + \alpha_{3k} f_k + \varepsilon_3 \\ \vdots \\ x_p = \alpha_{p1} f_1 + \alpha_{p2} f_2 + \alpha_{p3} f_3 + \alpha_{pk} f_k + \varepsilon_p \end{cases}$$

We use the matrix to express the mathematical model of factor analysis, such as:

$$X = AF + \varepsilon$$

F was named after factor, and A was named after loading matrix, αij (i=1, 2, …, p; j=1, 2, …, k) were named after factor loading.

The core of factor analysis is to solve the factor loading matrix. Solving methods are principal component analysis, least squares, maximum likelihood method, etc. We select principal component analysis to find the factor loading matrix. Principal component analysis transforms the original relevant variables xi which was standardized and linear combination into another unrelated variable yi, such as:

$$\begin{cases} y_1 = \mu_{11} x_1 + \mu_{12} x_2 + \mu_{13} x_3 + \cdots + \mu_{1p} x_p \\ y_2 = \mu_{21} x_1 + \mu_{22} x_2 + \mu_{23} x_3 + \cdots + \mu_{2p} x_p \\ y_3 = \mu_{31} x_1 + \mu_{32} x_2 + \mu_{33} x_3 + \cdots + \mu_{3p} x_p \\ \vdots \\ y_p = \mu_{p1} x_1 + \mu_{p2} x_2 + \mu_{p3} x_3 + \cdots + \mu_{pp} x_p \end{cases}$$

Where

$$\mu_{i1}^{\ 2} + \mu_{i2}^{\ 2} + \mu_{i2}^{\ 3} + \cdots + \mu_{ip}^{\ 2} = 1 (i = 1,2,3,\cdots, p)$$

The variable y1, y2, y3, …, yp were name after principal components of original x1, x2, x3, …, xp. Then we can find eigenvalue λ1≥λ2≥λ3≥⋯≥λp≥0 and eigenvector μ1, μ2, μ3, …, μp. With the eigenvalues and their corresponding eigenvector, we calculate the factor loading matrix:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p} \\ \vdots & \vdots & & \vdots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{p3} \end{pmatrix} = \begin{pmatrix} \mu_{11}\sqrt{\lambda_1} & \mu_{21}\sqrt{\lambda_2} & \cdots & \mu_{p1}\sqrt{\lambda_p} \\ \mu_{12}\sqrt{\lambda_1} & \mu_{22}\sqrt{\lambda_2} & \cdots & \mu_{p2}\sqrt{\lambda_p} \\ \vdots & \vdots & & \vdots \\ \mu_{1p}\sqrt{\lambda_1} & \mu_{2p}\sqrt{\lambda_2} & \cdots & \mu_{pp}\sqrt{\lambda_p} \end{pmatrix}$$

Because k<p, we choose the eigenvalues and their corresponding eigenvector, then we solve the factor loading matrix:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2k} \\ \vdots & \vdots & & \vdots \\ \alpha_{k1} & \alpha_{k2} & \cdots & \alpha_{kk} \end{pmatrix} = \begin{pmatrix} \mu_{11}\sqrt{\lambda_1} & \mu_{21}\sqrt{\lambda_2} & \cdots & \mu_{k1}\sqrt{\lambda_k} \\ \mu_{12}\sqrt{\lambda_1} & \mu_{22}\sqrt{\lambda_2} & \cdots & \mu_{k2}\sqrt{\lambda_k} \\ \vdots & \vdots & & \vdots \\ \mu_{1p}\sqrt{\lambda_1} & \mu_{2p}\sqrt{\lambda_2} & \cdots & \lambda_{kp}\sqrt{\lambda_k} \end{pmatrix}$$

Here we can find the factor loading.

Variance analysis models can be expressed like this (Wang and Sui, 2012)[7]:

$$\begin{cases} X_{ij} = \mu_j + \varepsilon_{ij} \\ \varepsilon_{ij} - N(0,\sigma^2) \\ i = 1,2,\cdots,n; \ j = 1,2,\cdots,s \end{cases}$$

$$\overline{X} = \frac{1}{n}\sum_{j=1}^{s}\sum_{i=1}^{n_j} X_{ij}$$

$$n = \sum_{j=1}^{s} n_j$$

$$SST = \sum_{j=1}^{s}\sum_{i=1}^{n_j}(X_{ij} - \overline{X})^2$$

$$\overline{X}_{.j} = \frac{1}{n}\sum_{i=1}^{n_j} X_{ij}$$

$$SSE = \sum_{j=1}^{s}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{.j})^2$$

$$SSA = \sum_{j=1}^{s}\sum_{i=1}^{n_j}(\overline{X}_{.j} - \overline{X})^2$$

$$SST = SSE + SSA$$

$$F = \frac{SSA/(s-1)}{SSE/(n-s)} - F(s-1, n-s)$$

The scales measured cultural landscape value were modified from (Song, 2008). The scales measured cultural benefits were modified from (Lv and Yu, 2010)[8]. The scales measured landscape select behavior of travelers were modified from (Hofstede, 2001)[2].

Liuzhou is historical and cultural city in China. Liuzhou was selected as cultural sample in this study. 350 travelers in Liuzhou formed the sample for the simulation. The investigation began in 12 March 2013 and ended in 12 April 2013. A total of 313 questionnaires were issued, finally 283 questionnaires were returned, 90.42% recovery rate.

To examine the reliability of the scales, we computed cronbach's alphas (Cronbachα) and construct reliability (CR) for the scales. The results were shown in TABLE 1 and TABLE 2. These values suggest a high internal consistency among the items and with their related constructs. To test the validity of scales, we test the discriminant validity, by conducted a confirmatory factor analysis and analyzed the covariance matrix using the maximum likelihood procedure of Amos 7.0. The results of discriminant validity were shown in TABLE 2, we compared the correlation coefficients between factors with the average variance extracted of the individual factors. This showed that the correlation coefficient between factors were lower than the average variance extracted of the individual factors, confirming discriminant validity. NV, OV, PV, CB, CLSBT refers to "national value", "organizational value", "personal value", "cultural benefits", "cultural landscape select behavior of travelers" respectively.

All the factor loadings were greater than 0.5, that means that there are three types cultural value exist which are national value, organizational value, personal value, and cultural benefits exist.

The factor loadings and R2, item-total correlation for national value (loading1=0.71, loading2=0.72, loading3=0.74; R12=0.70, R22=0.73, R32=0.72; correlation1=0.72, correlation2=0.73, correlation3=0.71); the factor loadings and R2, item-total correlation for organizational value (loading1=0.76, loading2=0.77, loading3=0.73; R12=0.76, R22=0.77, R32=0.76; correlation1=0.78, correlation2=0.76, correlation3=0.76); the factor loadings and R2, item-total correlation for personal value (loading1=0.72, loading2=0.72, loading3=0.71, loading4=0.77; R12=0.77, R22=0.76, R32=0.76, R42=0.76; correlation1=0.75, correlation2=0.71, correlation3=0.76, correlation4=0.76); the factor loadings and R2, item-total correlation for cultural benefits (loading1=0.74, loading2=0.74, loading3=0.77; R12=0.77, R22=0.73, R32=0.74; correlation1=0.71, correlation2=0.78, correlation3=0.78); the factor loadings and R2, item-total correlation for cultural landscape select behavior of travelers (loading1=0.75, loading2=0.76, loading3=0.76; R12=0.76, R22=0.74, R32=0.74; correlation1=0.76, correlation2=0.75, correlation3=0.76); those values show a good results of factor analysis.

RESULTS AND DISCUSSION

The variance analysis was conducted by Matlab 7.0. The results were shown in Figure 2, Figure 3, Figure 4 and Figure 5. Where NV, OV, PV, CB, CLSBT refers to "national value", "organizational value", "personal value", "cultural benefits", "cultural landscape select behavior of travelers" respectively.

National value and cultural benefits have variation phenomenon. Organizational value and cultural benefits have variation phenomenon. Personal value and cultural benefits have variation phenomenon. National value, organizational value, personal value and cultural benefits have variation phenomenon at the same time. Researchers aim at group behavior in bioinformatics can obtain insight from this study. Variation phenomenon is significantly in cultural group behavior.

TABLE 1 : Reliability test

	Cronbachα	CR
NV	0.80	0.81
OV	0.82	0.83
PV	0.81	0.82
CB	0.80	0.81
CLSBT	0.82	0.82

TABLE 2 : Discriminant validity test

NV	OV	PV	CB	CLSBT
0.77				
0.05	0.75			
0.07	0.04	0.73		
0.01	0.02	0.03	0.77	
0.01	0.01	0.01	0.06	0.78

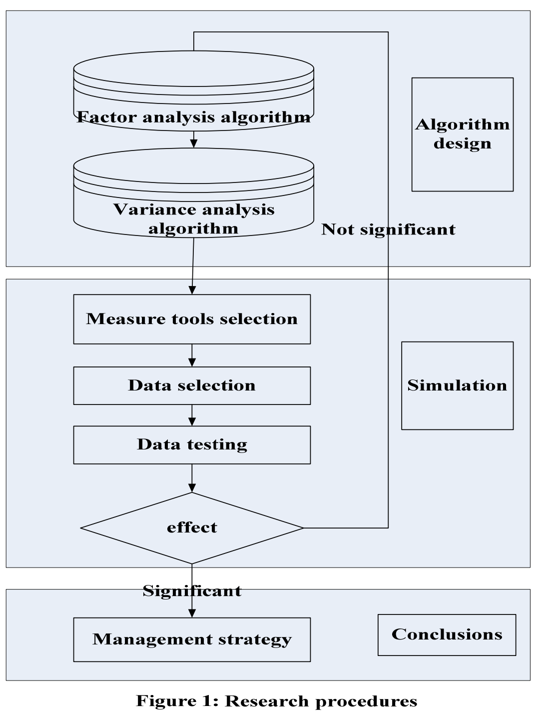

Figure 1 : Research procedures

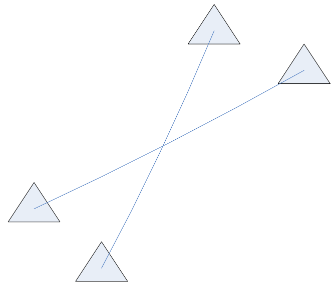

Figure 2: Variation phenomenon of national value and cultural benefits

Figure 2 : Variation phenomenon of national value and cultural benefits

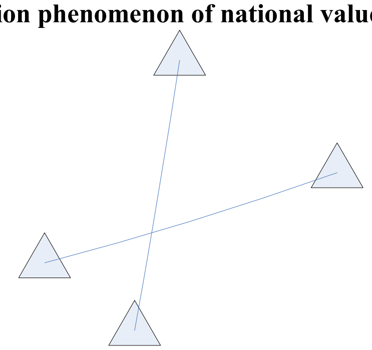

Figure 3: Variation phenomenon of organizational value and cultural benefits

Figure 3 : Variation phenomenon of organization value and cultural benefits

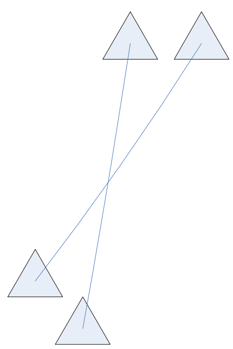

Figure 4: Variation phenomenon of personal value and cultural benefits

Figure 4 : Variation phenomenon of personal value and cultural benefits

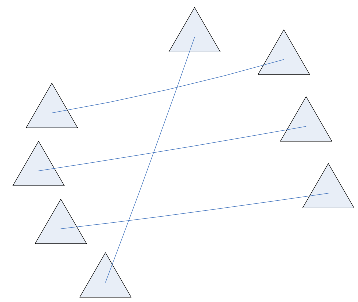

Figure 5: Variation phenomenon of cultural landscape value and cultural benefits

Figure 5 : Variation phenomenon of cultural landscape value and cultural benefits

ACKNOWLEDGMENTS

The authors wish to thank the helpful comments and suggestions from anonymous referee. The project sponsored by the Humanities and Social Science Research Project of Institutions of Higher Learning in Guangxi (SK13LX022).

REFERENCES

[1] K.Song, (Ed.); The Influence of Chinese Cultural Values on Consumer Perceptions and Behavioral Intention towards an Apparel Mass Customization Website. Iowa State University, **(2008)**.

[2] G.Hofstede, (Ed.): Culture’s Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations. Thousand oaks, 2001.

[3] Anderson, (Ed); The Economic Potential of Diversity: Creative Linkages between Internal and Market-Oriented Corporate Strategy. Fielding Graduate Institute, **(2002)**.

[4] Smith, (Ed.); Relationship Management in the Sales Organization: An Examination of Leadership Style and Cultural Orientation in Sales Manager and Salesperson Dyads. Drexel University, **(2004)**.

[5] W.Xue, (Ed.); SPSS Data Analysis. China Renmin University Press, **(2006)**.

[6] H.Harman, (Ed.); Modern Factor Analysis. University of Chicago Press, **(1967)**.

[7] Y.Wang, S.Sui, (Ed); Test Design and Matlab Data Analysis. Tsinghua University Press, **(2012)**.

[8] J.Lv, Z.Yu; Cultural benefits trial - based on the analysis of the perspective of economics. J.Fudan.Univers., **2**, 127-132 **(2010)**.