ISSN : 0974 - 7435 Volume 8 Issue 6

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 8(6), 2013 [750-756]

Study on the regression model of the relations between special quality and special results for high jump athletes

Yan Feng

Department of Physical Education, Zhengzhou University, Zhengzhou 450000, (CHINA)

E-mail: zb58@163.com

ABSTRACT

In order to study the relationship between the high jumpers' special quality and its results, this article collects 14 world-class excellent jump athletes' special quality and special performance on seven special items, as 100m run, standing triple jump, the run-up hand tough high, 4-6 strides approach high, after throwing the shot put, high grab the barbell and squat barbell. First of all, conduct correlation analysis of each special quality and its special performance. And the study found that there is certain correlation between the two. Secondly, conduct a principal component analysis of all special qualities and extract a principal component. Then build a regression model with the principal component data as independent variables and its special results as the dependent variables. This model reflects the relationship between the special qualities and special performances, and provides a reference for the daily training for coaches and athletes.

KEYWORDS

High jump;
Special quality;
Special performance;
Regression model.

INTRODUCTION

The high jump is a highly technical track and field event. With the development of the world track and field career, high jumpers have a very high level in their day-to-day training process and a growing number of athletes and coaches have started to pay attention to the training of special qualities. Therefore, in-depth study on the relationship of long jumpers' various special qualities and corresponding performances possesses great realistic significance.

Relative experts and scholars from home and abroad have conducted a lot of research work on the relationship between special performance and special quality. For example, Laomin Li's study shows that by strengthening high jump athlete's training on special qualities as snatching, squatting, triple jumping, jumper's relative special performance can be improved. Faping Chen, through statistical analysis of special achievements data and special qualities data and sequencing of long jumpers' eight special qualities as: run between 30 meters, the 100 meters run and the standing triple jump, finds that factors as standing triple jump, run-up jump high and the run-up hand tough height have the greatest impact on athletes' special performances. And previous studies are usually focused on the all various special qualities. Whereas, special qualities affecting high jump athletes' special performances are of many kinds, and it is impossible to take into account all the special qualities by the High Jumpers in the day-to-day training. Therefore, identifying several major aspects affecting the high Jumpers' special performances with an appropriate method, in order to further guide the training of coaches and athletes, is more realistic.

This study aims at conducting principal component analysis of various special qualities, achieving the drop-dimensional of special qualities. Then build a regression model with the main principal components and special achievements and analyze the importance degree of each principal component to special performance, in order to provide certain theoretical basis for high jumpers' daily training.

BRIEF INTRODUCTION OF THE METHOD

Correlation analysis

The correlation analysis is a statistical analysis method that studies the correlation between the random variables, mainly by the correlation coefficient to determine the correlation. For two variables, x and y, if the sample values are respectively x_i and y_i, then the correlation coefficient of the two is:

$$r_{xy}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}\in[-1,1]$$

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

If the absolute value of r_{xy} is closer to 1, it demonstrates that the closer is the relationship between the two variables; and if the absolute value is closer to 0, it indicates that the relationship between the two variables is less closely.

Principal component analysis

The basic principle of the Principal Component Analysis:

Suppose that the number of indexes is p, which are X1, X2,..., Xp. To find a set of independent composite indicators Z1, Z2,..., Zp that summarize the main information of the number p indexes, is to find a set of constants ai1,ai2,…,aip (i=1,2,…,p), that can realize the linear combination of the number p indexes, from a mathematical point of view:

$$\begin{cases} Z_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ Z_2 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{2p}X_p \\ \vdots \\ Z_p = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p \end{cases}$$

The linear combination can generalize the main information of the number p original indexes. For the convenience of description, introduce the following forms of matrix:

$$Z=\begin{Bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{Bmatrix},\quad A=\begin{Bmatrix} a_{11} & a_{12} & \cdots a_{1p} \\ a_{21} & a_{22} & \cdots a_{2p} \\ \vdots & & \\ a_{p1} & a_{p12} & \cdots a_{pp} \end{Bmatrix} \triangleq \begin{pmatrix} a_1^{'} \\ a_2^{'} \\ \vdots \\ a_p^{'} \end{pmatrix},$$

$$X=\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}$$

Then formula (1-1) can be expressed as Z=AX or

$$\begin{cases} Z_1 = a_1^{'}X \\ Z_2 = a_2^{'}X \\ \vdots \\ Z_p = a_p^{'}X \end{cases}$$

If formula $Z_1=a_1^{'}X$ satisfies $a_1^{'}a_1=1$ and $Var(Z_1)=\underset{a_1^{'}a_1=1}{Max}\{Var(a^{'}X)\}$, then Z1 is named as the first principal component of original indexes X1,X2,…,Xp. When Zi‴Zj, principal components Zi and Zj are unrelated, and Z1 is the linear combination with the largest variance among all the linear combination of original indexes X1, X2,…,Xp; Z2 is the linear combination with the largest variance among all the linear combination of original indexes X1,X2,…,Xp except for Z1. Sequence the indicators

FULL PAPER

according to the variance size, and so on.

Basic steps of principal component analysis of the basic steps:

1) Standardization of the raw data. With application of Z-score method, the transformation formula is:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

In the formula, $\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$,

$s_j = [\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2]$ i=1, 2, 3,…, n j=1,2,3,…,p

The mean of the transformed data is zero and the variance is 1.

The reason for original data standardization: in the principal component analysis, when determine the principal components by correlation coefficient matrix R (covariance matrix Σ), variables with larger variances δ_j^2 are often given priority to for consideration. In other words, this process is highly affected by the measurement scales of variables. Sometimes this will cause unreasonable results. In order to more objectively explain the connotation of the principal components, standardization of the original variable data must be conducted, avoiding the impact of the measurement units and magnitudes.

2) Determine the correlation matrix R of index data

The correlation coefficient matrix R of Variables (indicators) is the starting point of the principal component analysis. Measurement formula is:

$$r_{ik} = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(x_{ij} - \bar{x}_i)}{s_j}\frac{(x_{ik} - \bar{x}_k)}{s_k}$$ Or

$$r_{ik} = \frac{1}{n-1}\sum_{i=1}^{n} Z_{ij}Z_{jk}$$

And there is Rii=1and rik=rkj.

3) Determine the Eigen value, Eigen vectors and the contribution rate of matrix R

The characteristic equation of matrix R $|\lambda I_P - R|=0$, and λ_g (g=1, 2,…,P) is the characteristic root of the characteristic equation. It is the variance of principal component Z and its value describes each principal component‘s ability to summarize the original indexes' information. And L represents a P-dimensional real vector. The vector Lg calculated from equations $[\lambda_g I_P - R]L_g=0$ is the corresponding characteristic vector to characteristic root λ_g. It is also the coefficient of each component on the new coordinates of standardized vector. $Z_j\left[Z_i = \begin{pmatrix} z_{1j} \\ \vdots \\ z_{pj} \end{pmatrix}\right]$ ag= $\lambda_g / \sum_{g=1}^{p} \lambda_g$, this formula shows each principal component explains the information of the original variables, that is variance contribution rate.

Stepwise regression analysis

Regression analysis is used to study the relationship between objective things. Built on the basis of a large number of observations and test of objective things, it is a statistical method to find those statistical regularity hidden in the outlook of uncertain phenomenon. In terms of theory and practice, multiple linear regression technique is a fairly classic quantitative analysis method.

The research objects of multiple linear regression model is affected by multiple factors: $X_1, X_2, \cdots, X_n$. Aassume that the relationship of various factors and Y is linear and then the multiple linear regression models are:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i$$

Wherein, $\beta_0, \beta_1, \cdots, \beta_p$ is the numberunknown parameters, $\beta_1, \cdots, \beta_p$ is called as the regression constant, β_0 called as the regression coefficient, y is known as the dependent variable, $x_1, x_2, \cdots, x_n$ is the general variables that can accurately measured and controlled, also called independent variables, ε isRandom error. There is the following hypothesis about the random error:

$$\begin{cases} D(\varepsilon_i) = \sigma^2 & \\ E(\varepsilon_i) = 0 & i = 1, 2, \cdots, n \\ \operatorname{cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} \sigma^2, i = j \\ 0, i \neq 0 \end{cases} & i, j = 1, 2, \cdots, n \end{cases}$$

To a practical problem, if the set number of observed data that can be obtained is n: $(x_{i1}, x_{i2} \ldots, x_{in}; y_i)$ and $i = 1,2,\ldots,n$. Then the linear regression model formula $(4-20)$ can be expressed as:

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} + \varepsilon_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_p x_{2p} + \varepsilon_2 \\ \cdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_p x_{np} + \varepsilon_n \end{cases}$$

In the form of matrix is $y = XB + \varepsilon$

Wherein:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & & \cdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, B = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_p \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{bmatrix}$$

Estimate the regression parameters:

Estimate with the least squares method. Suppose the residual value between the true value and the estimated value of the model, then:

Wherein: $\hat{Y} = XB$, $\hat{Y} = XB$

According to the requirements of the least squares method, there should be: $E'E = (Y-\hat{Y})'(Y-\hat{Y}) = \text{Min}$, $E'E = (Y-XB)'(Y-XB) = \min$

Pole value principle: according to differential coefficient of matrix, determine the derivative of matrix B, and suppose the value is 0, then:

$$\frac{\partial E'E}{\partial B} = \frac{\partial (Y-XB)'(Y-XB)}{\partial B} = \frac{\partial (Y'Y - 2Y'XB + B'X'XB)}{\partial B} = -2(Y'B)' + 2(X'X)B = 0$$

The estimated value of the regression coefficient vector, B is: $\hat{B} = (X'X)^{-1} X'Y$

In accordance with the application conditions of the model:

(1) The dependent and independent variables have a linear relationship;

(2) The observations, $Y_i (i = 1,2,\cdots,n)$, are independent of each other;

(3) The residual error e obeys the normal distribution with mean 0 and variance σ^2.

Select the training sample and test sample of the BP neural network modeling for modeling, and simulation samples are used to predict the multiple linear regression modeling. The predictive analysis sequence is modeling, inspection and prediction.

STEPWISE REGRESSION MODEL OF THE RELATIONSHIP BETWEEN HIGH JUMPERS' SPECIAL QUALITY AND SPECIAL PERFORMANCE

Research data

14 world elite male high jumpers' data of specific performance and the special quality is collected, as shown in TABLE 1.

Descriptive analysis of the 14 athletes' special qualities and special performances was performed, and the results are shown in TABLE 2.

TABLE 1 : 14 world elite male high jumpers' data of specific performance and the special quality

Number	100m run /s	Standing triple jump /m	Run-up hand touch height /m	4-6 steps run-up height /m	After throwing shot /m	High grip barbell /m	Squat barbell /m	Specific Performance /m
1	10.7	10.1	1.25	2.25	16	135	200	2.4
2	10.8	10.06	1.24	2.24	15.9	132.5	197.5	2.39
3	10.8	10.01	1.23	2.23	15.8	132.5	195	2.38
4	10.8	9.95	1.22	2.22	15.7	130	192.5	2.37
5	10.8	9.92	1.21	2.21	15.6	130	190	2.36
6	10.8	9.87	1.2	2.2	15.5	127.5	187.5	2.35
7	10.9	9.85	1.19	2.2	15.4	127.5	185	2.34
8	10.9	9.75	1.19	2.19	15.3	127.5	185	2.33
9	10.9	9.7	1.18	2.18	15.2	125	180	2.32
10	10.9	9.63	1.18	2.18	15	125	177.5	2.31
11	10.9	9.55	1.17	2.16	14.6	120	172.5	2.3
12	10.9	9.55	1.16	2.16	14.6	120	172.5	2.29
13	11	9.52	1.16	2.16	14.4	120	170	2.28
14	11	9.45	1.16	2.15	14.2	117.5	170	2.27

FULL PAPER

TABLE 2 : The descriptive analysis result of the 14 athletes' special qualities and special achievements

Statistics	100m run /s	Standing triple jump /m	Run-up hand touch height /m	4-6 steps run-up height /m	After throwing shot /m	High grip barbell /m	Squat barbell /m	Specific Performance /m
Mean	10.8643	9.7793	1.1957	2.195	15.2286	126.4286	183.9286	2.335
Standard deviation	0.08419	0.21614	0.03031	0.03205	0.58366	5.43442	10.36425	0.04183
Maximum	10.7	9.45	1.16	2.15	14.2	117.5	170	2.27
Minimum	11	10.1	1.25	2.25	16	135	200	2.4

Correlation analysis of special quality and specific performance

With person correlation analysis, determine the correlation coefficient of each special quality and the specific performance, and the results is in TABLE 3.

As can be seen from TABLE 3, there is correlation between each special quality and its performance.

Principal component analysis of the special qualities

The seven special qualities, 100m run, standing triple jump, the run-up hand tough height, 4-6 strides approach height, after throwing the shot, the high grip barbell and squat barbell are expressed as $x_1, x_2, \cdots, x_7$ respectively. Conduct principal component analysis of the seven variables, and partial correlation coefficients among variables are shown in TABLE 4.

Conduct significance test of each partial correlation coefficient, and P is less than 0.05, indicating that there are correlation between variables. Therefore it is necessary to conduct the principal component analysis.

TABLE 3 : Correlation analysis results between the special quality and the special performance

Special performance /m	100m run /s	Standing triple jump /m	Run-up hand touch height /m	4-6 steps run-up height /m	After throwing shot /m	High grip barbell /m	Squat barbell /m
r	-0.928	0.995	0.983	0.990	0.983	0.981	0.994
P	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 4 : The statistical information of the principal component

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.745	96.360	96.360	6.745	96.360	96.360
2	.154	2.202	98.562			
3	.069	.981	99.543			
4	.018	.252	99.795			
5	.009	.124	99.919			
6	.004	.058	99.978			
7	.002	.022	100.000			

The statistical information of the principal components is shown in TABLE 4. As can be seen from TABLE 4, the characteristic root of the first principal component is 96.36%, meaning the cumulative contribution rate reaches to 96.36%. Thus only select a principal component. The gravel figure of principal component analysis result is shown in Figure 1.

Matrix of factor loadings of the principal component analysis is shown in TABLE 5. As can be seen from TABLE 5, the first principal component contains the information of all the variables.

Judging from the matrix of factor loadings, expressions of the first principal component can be:

$$z_1 = -0.932 * stdx_1 + 0.994 stdx_2 + 0.984 stdx_3 + 0.992 stdx_4 + 0.984 stdx_5 + 0.988 stdx_6 + 0.996 stdx_7$$

According to the above equation, the value of the first principal component can be obtained.

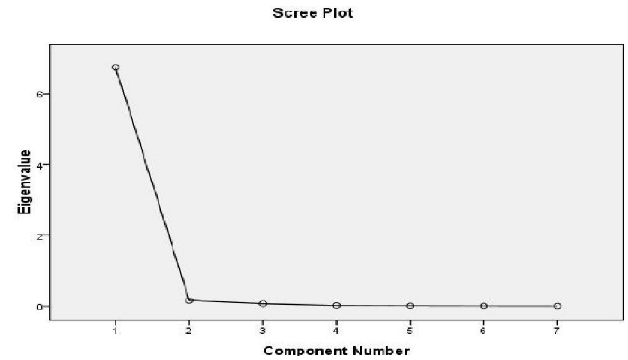

Figure 1 : Gravel figure of principal component analysis

Build the stepwise regression model

The first principal component as independent variables and high jumpers' special performance as dependent variables, conduct the regression analysis. The coefficient of the model is R^2 =0.995, indicating good fitting effect. The parameters of the model are shown in TABLE 6.

Judging from TABLE 6, the final established regression equation is:

$$y = 2.335 + 0.042z_1 = 2.335 + 0.042*(-0.932*stdx_1 + 0.994stdx_2 + 0.984stdx_3 + 0.992stdx_4 + 0.984stdx_5 + 0.988stdx_6 + 0.996stdx_7)$$

TABLE 5 : Matrix of factor loadings

	100m run /s	Standing triple jump /m	Run-up hand touch height /m	4-6 steps run-up height /m	After throwing shot /m	High grip barbell /m	Squat barbell /m
Component 1	-0.932	0.994	0.984	0.992	0.984	0.988	0.996

TABLE 6 : Parameter estimation results of the model

Parameter	Partial regression coefficient	Standardized partial Regression coefficient	*t*	*P*
a	2.335	-	2923.290	.000
b	0.042	0.998	50.348	.000

CONCLUSIONS

(1) In this study, the 100m run, standing triple jump, the run-up hand tough height, 4-6 strides approach high, after throwing shot, high grip barbell and squat barbell are closely related with the high Jumpers' special performance.

(2) Build the prediction model of elite high Jump athletes:

, respectively represents for the 100m run, standing triple jump, the run-up hand tough height, 4-6 stride approach height, after throwing shot, high grip barbell and barbell squat. This model reflects the relationship between high jump athletes' special performances and special qualities and provides certain reference for the day-to-day training of the athletes.

REFERENCES

[1] Baocheng Wang, Wenchao Li; The Related Analysis of the Specialized Quality and Results of Excellent Male High Jumpers in Beijing. Journal of Capital Institute of Physical Education, **18(2)**, 72-75 **(2006)**.

[2] Bing Zhang,Hui Yue; Bio-mechanical Mathematical Model Analysis for Race Walking Technique. International Journal of Applied Mathematics and Statistics, **40(14)**, 469-476 **(2013)**.

[3] Bing Zhang,Yan Feng; The Special Quality Evaluation of the Triple Jump and the Differential Equation Model of Long Jump Mechanics Based on Gray Correlation Analysis. International Journal of Applied Mathematics and Statistics, **40(10)**, 136-143 **(2013)**.

[4] Bing Zhang; Dynamics Mathematical Model and Prediction of Long Jump Athletes in Olympics. International Journal of Applied Mathematics and Statistics, **44(14)**, 422-430 **(2013)**.

[5] Brueggemann; 6th World Championships biomechanical study. Chinese Athletic Association, **4**, 22-24 **(1998)**.

[6] Cai Cui; Application of Mathematical Model for Simulation of 100-Meter Race. International Journal of Applied Mathematics and Statistics, **42(12)**, 309-316 **(2013)**.

[7] Faping Chen; Analyzing on Gray Correlatives between the Body Quality of the Top High jumpers in the World and Performance of the High jumpers. Journal of Guangzhou Physical Education Institute, **23(3)**, 57-60 **(2003)**.

[8] Gang Lu, Yongxiang Qin; An Analysis of the Advantages of Fosbury Flop: A Biomechanics of View. Sports & Science, **19(4)**, 63-64 **(1998)**.

[9] Haibin Wang, Shuye Yang; An Analysis of Hurdle Performance Prediction Based On Mechanical Analysis and Gray Prediction Model. International Journal of Applied Mathematics and Statistics,

39(9), 243-250 **(2013)**.

[10] Hongwei Yang; Evaluation Model of Physical Fitness of Young Tennis Athletes Based On AHP-TOPSIS Comprehensive Evaluation. International Journal of Applied Mathematics and Statistics, **39(9)**, 188-195 **(2013)**.

[11] Hui Han; A new understanding of the high jump techniques. Chinese Athletic Association, 6, 17-20.

[12] Jiajin Liu, 2003,Study of GM(1,h) Model in Track and Field Training. Journal of Xi'an Institute of Physical Education, **20(6)**, 57-58 **(1997)**.

[13] Jiemin Chen, Ying Zhang; Kinematic Analysis on Take-off Technique of Chinese Elite Male High Jumpers. China Sport Science and Technology, **4**, 37-40 **(2001)**.

[14] Lei Yuan; Construction of Recurrent Neural Network Model for Relationship between Specific Performance and Physical Fitness of Top High Jumpers in the World. Journal of Beijing Sport University, **31(2)**, 202-204 **(2008)**.

[15] Qinhai Fan, Yue Zhou, Jian Zhou; The Research on the Relationship between Physical Fitness and Performance of High Jumpers. China Sport Science and Technology, **38(2)**, 35-38 **(2002)**.

[16] Rixin Hou; Biomechanical Analysis and Technical Characteristics of Approach Run in Arc in Fosbury Flop. Journal of Beijing University of Physical Education, **22(2)**, 59-61 **(1999)**.

[17] Yi Liu; The Establishment of Hierarchical Model for Basketball Defensive Quality. International Journal of Applied Mathematics and Statistics, **44(14)**, 245-252 **(2013)**.

[18] Yong Fan; Statistical Analysis Based On Gray System Theory Basketball Team Scores Its Technical Indicators Associated. International Journal of Applied Mathematics and Statistics, **44(14)**, 185-192 **(2013)**.

[19] Yuexian Song, Lei Yuan; Kinematics Analysis of Zhang Li-wen's High Jump Technique. Sichuan Sports Science, **3**, 28-30 **(2002)**.

[20] Zuojun Tan; Fuzzy Data Envelopment Analysis and Neural Network Evaluation Mathematical Applications Model Which Based On Martial Arts Competition. International Journal of Applied Mathematics and Statistics, **44(14)**, 37-44 **(2013)**.