

Establishment of Multivariate Statistical Regression Model Based on Technical Level and Cultivation Influencing Factors and Its Application in the Research of Basketball Reserve Talents

Sha Ji

College of Sports Science, Harbin Normal University, Harbin, China
ji_sh_a@126.com

Keywords: basketball reserve personnel; technical level, regression analysis; CUBA

Abstract. After the statistic analysis and the investigation of the 14th CUBA qualifying event, the problems of the competitiveness of the CUBA players and the basketball reserve personnel training have been discussed in this article. And the relation between the technical specifications and the achievements of CUBA teams have been analyzed by the stepwise multiple regression method. Nine technical specifications have been adapted in this article, such as 2-point shooting, 3-point shooting, free throw percentage, rebound, assist, foul, miss, steal, block, as the independent variables. The scores got and lost by each team in each event are taken as the dependent variables. Then the optimal regression equations of the scores got and lost by 12 CUBA teams and 9 technical specifications have been established by using the stepwise regression method. The relative coefficients represent the each technical specification's contribution to the achievements of the teams. The average of each team's technical specifications and the proportion of scores lost and got have been calculated. The relative coefficient r of each technical specification and the proportion of scores lost and got is got by SAS program. According to the value of r , we can get the sequence of the technical specifications' contribution to the achievements. The sequence from large to small is 2-point shooting, assist, steal, rebound, free throw percentage, block, 3-point shooting, miss, and foul. Therefore, it can provide technical support to the cultivation of reserve personnel in our country.

Introduction

In the past decades, the basketball talents were trained since the primary school, then entering the sports school and professional team. At that time, the professional teams in planning economy attracted many teenagers and many parents believed that sending children to the army, province team, or city professional team was an ideal way. The material guarantee was the main reason of the relative stability of the basketball talents. But the children trained in the professional teams lacked the cultural education, the communication with the society, and the scientific training. So after several years, what they had would be only the illness and uncertain future. So we need a new system to replace it, which is also the reason of the birth of CUBA..

The development of the college basketball activities promotes the popularity of the basketball in primary and middle schools. More teenagers take part in the basketball and then the counterpart personnel training system of primary school, middle school and college have been formed. The pyramid of Chinese basketball talents is established.

After seeing the CUBA competitions, many people think the professional teams are better. Indeed, the college players are immature both in physical qualifications and technical levels. In this article, by analyzing the different CUBA teams and performance of the players, we find a better solution for the basketball reserve personnel training in our country.

The Analysis of Technical Specifications

A. *The Analysis of Regression*

According to the analysis of the technical index, the main factors influencing the scores are the following technical index. x_1 is 2-point shooting; x_2 is 3-point shooting; x_3 is free throw percentage;

x_4 is rebounds; x_5 is the number of times of assist; x_6 is the number of times of foul; x_7 is the number of times of miss; x_8 is the number of times of steal; x_9 is the number of times of block. Considering there is ability difference among teams, we regard the proportion of scores got and lost of each team Z_j as the achievement index.

If the geometrical shape of curve of index is more similar to the curve of proportion of scores got and lost and the changes of slope are more same, the relationship is closer. From the table, we can say that 2-point shooting and free throw percentage have closer relationship with the scoring rate. But in this way can we see the similar curves and get some indexes which have closer relationship with the scoring rate. We can't get the relation between all indexes and the achievements and can't get detailed relative data.

Therefore, we regard the score got and lost for each team Z_j as the dependent variable and regard the nine technical indexes as the independent variables. Then we choose the important variables from the many technical indexes by the stepwise regression method and establish the optimal regression equation $Z_i = f(x_i)$ of the proportion of scores got and lost Z_i and the technical indexes x_i . This regression equation includes the technical indexes both significantly affecting the achievements and not.

Each team attends 5 competitions. We make the stepwise regression of the competition data and the proportion of scores lost and got by SAS program. Then the regression equations of each team are got as below,

$$\text{CUBA Team One: } z = 0.66668 + 0.03677x_5 ;$$

$$\text{CUBA Team Two: } z = 0.25383 - 0.71245x_1 + 0.07244x_5 + 0.10382x_8 ;$$

$$\text{CUBA Team Three: } z = 0.25383 - 0.71245x_1 + 0.07244x_5 + 0.10382x_8 ;$$

$$\text{CUBA Team Four: } z = 0.3888 + 0.57416x_1 + 0.01505x_5 + 0.00857x_7 ;$$

$$\text{CUBA Team Five: } z = 1.39221 + 0.11454x_1 + 0.75772x_3 + 0.00298x_7$$

$$\text{CUBA Team Six: } z = 0.69296 + 0.12114x_9 ;$$

$$\text{CUBA Team Seven: } z = 0.64396 + 0.03795x_5 - 0.07209x_9 ;$$

$$\text{CUBA Team Eight: } z = 0.5735 + 2.15444x_1 - 0.83891x_3 ;$$

$$\text{CUBA Team Nine: } z = 0.87614 + 0.02029x_4 - 0.01282x_8 - 0.15433x_9 ;$$

$$\text{CUBA Team Ten: } z = 1.52566x_1 + 0.00730x_4 ;$$

$$\text{CUBA Team Eleven: } z = 0.09003x_5 + 0.02481x_6 - 0.43080 ;$$

$$\text{CUBA Team Twelve: } z = 3.34512x_3 - 0.09664x_8 ;$$

From the above regression equations, we can clearly see that different teams have different technical indexes which can affect the achievements.

B. *The Relative Coefficient Analysis*

In order to assure the influence of the indexes on the team achievements, we use correlation analysis method to calculate the relative coefficient and evaluate the indexes contribution to the achievements according to the relative coefficients. Because each team attends 5 contests, we get 60 groups statistic data. Due to the large amount of data, we just use average of technical statistics to calculate. The followings are some averages of the each team.

TABLE I. SOME AVERAGE INDEX VALUES OF EACH FACULTY

Team	2-Point Shooting	3-Point Shooting	Free Flow Percentage	Block	Scoring Rate
A	0.50	0.35	0.68	1.4	0.96372
B	0.53	0.40	0.69		2.2	0.97246
C	0.52	0.39	0.66		2.2	0.95948
D	0.38	0.34	0.57		2.8	0.7685
E	0.54	0.27	0.64		4.4	1.12824
F	0.48	0.35	0.71		0.8	1.00416
G	0.58	0.44	0.71		3.4	1.13356
H	0.51	0.40	0.77		2.6	1.12282
I	0.39	0.27	0.71		1	0.82125
J	0.54	0.40	0.68		2.6	1.20424
K	0.56	0.29	0.71		2.8	1.17358
L	0.55	0.34	0.70		2.7	1.18891

The two groups of random variables x_i and Z have been tested for 12 times randomly and the observed values are respectively $(x_i^{(j)}, Z_j)$

$$\bar{x}_i = \frac{1}{12} \sum_{j=1}^{12} x_i^{(j)}, \quad \bar{Z} = \frac{1}{12} \sum_{j=1}^{12} Z_j$$

The related coefficient of random variables x_i and z for the sample $(x_i^{(j)}, Z_j)$, $i=1,2,3 \dots 9$, $j=1,2, \dots, 11,12$ is as below,

$$R_i = \frac{\sum_{j=1}^{12} (x_i^{(j)} - \bar{x}_i)(Z_j - \bar{Z})}{\sqrt{\sum_{j=1}^{12} (x_i^{(j)} - \bar{x}_i)^2 \sum_{j=1}^{12} (Z_j - \bar{Z})^2}}$$

R_i can be regarded as the estimated value of the related coefficient r of the i^{th} index and the rate of score got and lost.

Use the related coefficient r to represent the degree of correlation of nine technical indexes to the proportion of scores got and lost. Then we can get the contribution of each index to the achievements.

After analyzing the data of the twelve teams by SAS program, we can get the related coefficient r of each technical index and the proportion of the scores got and lost. When r is bigger which shows the more contribution of the technical indexes to the achievements, the value can provide more impact on the result of the competitions. Otherwise, the impact will be less.

TABLE II. THE TECHNICAL INDEXES OF EACH TEAM AND GAIN AND LOSS

Indexes	The Related Coefficient
2-Point Shooting	0.8681
Assist	0.6249
Steal	0.6159
Rebound	0.5278
Free Throw Percentage	0.497
Block	0.4856
3-Point Shooting	0.2592
Miss	0.1315
Foul	0.0039

According to the value of r , we can see the contribution sequencing as follows, 2-point shooting, assist, steal, rebound, free throw percentage, block, 3-point shooting, miss and foul.

The Analysis of the Influencing Factors

The development of the CUBA reserve personnel can be influenced by many factors. After the above analysis, we determine the following 12 factors in Table 3 as the objectives of research.

TABLE III. THE FACTORS INFLUENCING THE DEVELOPMENT OF CUBA RESERVE PERSONNEL

Marks	Factors	Average	Sequence
A	Policy Guidance	4.18	7
B	Scientific Management	4.04	10
C	The Proficiency of Coaches	4.54	2
D	Competition System	4.43	3
E	External Support	4.07	9
F	Scientific Selection	4.68	1
G	Training System	4.36	4
H	Conditions of Cultivation	4.32	5
I	Personal Motivation	4.14	8
J	Environment of Basketball	4.00	11
K	Sports Science	4.21	6
L	The Competition of High-tech Talents	4.18	12

Table 4 is the common factor variance table. From this table, the variances of the common factors are all over 0.65. We can get Table 5 after statistically analyzing the initial characters of the 12 variables and cumulative contribution rate (0.69). Then we implement the relevance analysis of the 12 variables (shown in Table 6); we can find that both DAGH and KFBC are of correlation, so we find that dividing 12 variables into 3 common factors, which are government behavior factor, science and technology content factor and environmental factor, is applicable.

TABLE IV. THE COMMON FACTOR VARIANCE

	Before Selection	After Selection
<i>A</i>	1.00	0.83
<i>B</i>	1.00	0.89
<i>C</i>	1.00	0.78
<i>D</i>	1.00	0.79
<i>E</i>	1.00	0.78
<i>F</i>	1.00	0.79
<i>G</i>	1.00	0.79
<i>H</i>	1.00	0.81
<i>I</i>	1.00	0.68
<i>J</i>	1.00	0.82
<i>K</i>	1.00	0.71
<i>L</i>	1.00	0.79

TABLE V. COMMON FACTORS CLASSIFICATION AND FACTORS LOAD TABLE

Common Factors	Indexes	Influencing Factors	Factor Loadings
<i>Government Behavior Factors</i>	D	Competition System	0.59
	A	Policy Guidance	0.56
	G	Cultivation System	0.48
	H	Training Conditions	0.45
<i>Science and Technology Content Factors</i>	K	The Scientific Nature of Training	0.61
	F	Scientific Selection	0.51
	B	Scientific Management	0.49
	C	Proficiency of Coaches	0.43
<i>Environmental Factors</i>	J	Basketball Environment	0.56
	E	External Support	0.44
	I	Personal Motivation	0.44
	L	The Competition of High-tech Talents	0.42

TABLE VI. THE RELEVANCE ANALYSIS OF FACTORS

	A	B	C	D	E	F	G	H	I	J	K	L
A	1.00											
B	0.12	1.00										
C	0.21	.378*	1.00									
D	.375*	0.22	0.27	1.00								
E	0.06	0.16	0.31	0.13	1.00							
F	0.11	.430*	.368*	0.16	0.26	1.00						
G	.476*	0.06	0.21	.382*	0.16	0.30	1.00					
H	0.32	0.22	0.15	.421*	0.08	0.24	.367*	1.00				
I	0.19	0.18	0.28	0.23	.496*	0.11	0.24	0.16	1.00			
J	0.21	0.11	0.03	0.16	.432*	0.25	0.16	0.10	.396*	1.00		
K	0.18	.377*	.401*	0.15	0.30	.401*	0.24	0.18	.367*	0.26	1.00	
L	0.10	0.13	0.30	0.22	.372*	0.05	0.28	0.20	.451*	.401*	0.25	1.00

Conclusion and Suggestions

(1) At present our CUBA players don't perform well. They have problems in training, management and the competitions. The players think that they have long way to go to be in the line of the international excellent players. As the main reserve personnel training mode in China, CUBA has great potentials.

(2)The 12 elements influencing the proficiency of CUBA can be divided into 4 factors. The coaches should make the CUBA players improved in these aspects.

(3)Among these factors, the proficiency of the coaches and training conditions have the ultimate impacts on CUBA. The leaders' attention is also very important. The university should invest more in this aspect.

(4)The examination system should be reformed; we should relax the restriction of entrance to the students who are good at basketball.

Acknowledgement

We get the support from Humanities and Social Sciences Project of Education Department in Heilongjiang Province, which is called the 14th CUBA Qualifying Event: The Situation of the Basketball Players in Our Province and the Mode of Talents Cultivation. The series number of the project is 12522144.

References

- [1] Yu Jialin, Multivariate statistics and SAS Application [M], Wuhan, Wuhan University Press, 2008
- [2] Li Daqian, The Optimal Model of Press Resource Allocation [J], Journal of Engineering Mathematics, 2006
- [3] Chu Hongsheng, Instance Textbook of MATLAB7.2 Optimal Design [M], Beijing: Mechanical Engineering Press, 2001
- [4] Song Xiaochun, The Analysis of Rivalry Characteristics in Basketball Competition [J], Wuhan Institute of Physical Education Journal, 2010,34(6):50-51. [J]
- [5] Shen Lianguo, Xu Chenggang, The Approach and Strategy of Cultivating the High Level Competitive Players in Sports Colleges [J], 2000,34(6):26-28.
- [6] Bai Xilin and Huang Wenyan, The Continuous Development of Chinese Basketball and Basketball Reserve Personnel [J], Beijing Sport University, 1999
- [7] Xia Wen and Ling Ping, The Analysis of School Running Status of Teenagers Amateur Sport School in Zhejiang Province [J], Zhejiang Sports Science, 2010
- [8] Zhang Peifeng and Zhang Wei, The Comparison of the Influencing Factors of China-USA College Basketball Reserve Personnel Cultivation [J], Chengdu Institute of Physical Education, 2007