

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(14), 2014 [7689-7694]

The application of principal component analysis and cluster analysis in comprehensive evaluation for the NBA

Ke wang

Shandong Sport University, Rizhao, 276826, (CHINA)

ABSTRACT

This research mainly through to the 2013 season in the NBA basketball game in his first season in pre-season technical data analysis of each project. By using the statistical methods of principal components analysis for each project, the multiple statistical index clustering scores for binary ball shot composition, binary composition, free-throw scores composition, composition of shots, and the error component in the game. At the same time according to each component to the importance of the impact on the overall variables combined with the characteristics of the NBA teams put forward a few research, the game players dribbling, passing, and widen the vision control ball skill training is of great significance to reduce mistakes and other technical points. Through the above analysis of the NBA basketball game in each link, using principal component analysis and clustering analysis, and applies this method to the comprehensive evaluation of the NBA basketball game. Principal component analysis (pca) and cluster analysis method for the NBA basketball game is very extensive in application of comprehensive evaluation, the research for improving our country on the principal component analysis (pca) and cluster analysis method research has the very vital significance.

KEYWORDS

Principal component analysis; Clustering; Basketball.



INTRODUCTION

The player's comprehensive quality. Because we studied the technical points are a matter of a player's overall play as we know, NBA basketball game is the current international is the highest level of professional basketball game, in the process of the NBA game and mobilize the measure index is very much, these indicators mainly want to reflect a and the development of the whole team. At the same time, the technical points are linked to the achievement of the game. It is important to note that in these statistical data, what data is has direct relationship to the game, or about which data can be directly the result of a game, it is very critical. We also study how these statistics for effective statistics and analysis and distinguish between. This study selected in each of the players in the NBA some relatively typical data. According to the method of principal component analysis, find out the main components of the statistical data, and then to clustering of statistical indicators, according to the statistical data clustering analysis, find out the trend of the data, and show the general development direction of the current international basketball, by understanding the general direction of international basketball, can put forward the development of basketball in China has guiding significance to the development target and development direction, we can foster strengths and circumvent weaknesses, to find their own advantages and disadvantages, put forward the coping strategies of problems of the current development of basketball in China. The progress of China's basketball career, how to be more scientific and effective to daily training for the national men's basketball and women's basketball team and before the tour has a very high guidance role. To guide the future training and competition of Chinese men's basketball team, not only to the Chinese men's basketball team improve training pertinence and effect of the game, especially for Chinese youth basketball athletes training and competition have profound significance.

THE OBJECT OF STUDY

The 2013 season NBA preseason played basketball league of eastern and western part of the team going into the playoffs technical data statistics.

The research method

Mathematical statistics: Use SPSS10.0 statistical software for every team in the NBA to principal component analysis, statistical indicators data of these statistical data to carry on the principal component analysis, and then carry on the clustering to determine the number of components, the last step to determine the composition of the typical within the composition.

Logical analysis

Adopt the method of logical analysis, the data for a certain classification, according to our data for the NBA preseason game statistics, direction for the future development of basketball in our country according to principal component to grasp the development trend of world basketball, find out the highest level of our country and international basketball the deficiencies and gaps between the NBA.

THE RESULTS AND ANALYSIS

The principal component analysis

Principal component analysis is applied in many fields is more a kind of analysis method, a lot of times. In the process of study, we in order to more comprehensive and objective to study the cause of a problem, and each link, need to look at problems from the perspective of many. In turn, to analyze problems, considering the main object, we study and then to analyze the distribution, the principal component analysis is we often used in the analysis method. On the other hand, when we in the study of any one thing. We think of when to go to a certain point in discussing and studying, hard to avoid can appear a certain deviation, or incomplete, one-sided view and not scientific research results, this is not conducive to the study. The same research also can't get a very comprehensive scientific research results, finally the conclusion is not correct, let alone bring any guide for follow-up studies. If a separate analysis of these indicators, will no doubt cause the one-sided understanding of the research object. It is not easy to comprehensive and effective conclusions. Principal component analysis is to consider the relationship of each variable, using the ideas of dimension reduction of multiple indexes into less several unrelated comprehensive index, to make further study on simple statistical method. Reflect the same characteristics of variable has a lot of things. But to find each other independent and representative of the independent variable, without loss of most of the information.

Component analysis steps and results in the SPSS10.0 to establish database, click Analyze Data Reduction Factor. Finally concluded that the results are shown in TABLE 1, through the analysis of the composition, based on the principle of composition eigenvalues greater than 1, or composition of the principle of the cumulative percentage of the total variance is more than 80%, to determine the composition of the extraction of five major, we from the data given in the TABLE below, and the data reflected problems.

TABLE 1 : The principal factor analysis table

ingredient	Correlation matrix eigenvalues			Without rotation factor loading of sum of squares			After rotating the factor loading of sum of squares		
	Characteristic value of each component	Factor accounted for the proportion of the total variance	accumulative total %	Characteristic value of each component	Factor accounted for the proportion of the total variance	accumulative total %	Characteristic value of each component	Factor accounted for the proportion of the total variance	accumulative total %
1	6.341	31.703	31.703	6.341	31.703	31.703	4.617	23.084	23.084
2	4.715	23.575	55.278	4.715	23.575	55.278	4.590	22.948	46.032
3	2.600	12.999	68.277	2.600	12.999	68.277	4.307	21.535	67.567
4	1.846	9.231	77.508	1.846	9.231	77.508	1.842	9.210	76.777
5	1.263	6.316	83.824	1.263	6.316	83.824	1.409	7.047	83.824
6	0.851	4.257	88.081						
7	0.725	3.625	91.705						
8	0.588	2.942	94.674						
9	0.354		96.416						
10	0.306		97.947						
11	0.183		98.861						
12	0.125		99.485						
13	6.343E-02		99.802						
14	3.202E-02		99.962						
15	4.843E-03		99.986						
16	1.417E-03		99.993						
17	8.358E-04		99.997						
18	5.285E-04		100						
19	1.317E-05		100						
20	3.807E-06		100						

Clustering analysis

Variable clustering is also known as R type cluster in statistics, statistics of multi-index by computing the correlation coefficient between every index and clustering coefficient. According to the principle of "birds of a feather flock together", the nature and properties of indicators of similar kind of analysis method. The clustering results is shown as TABLE 2.

TABLE 2 : The clustering results

Points Per Game	Binary ball		three-pointer				penalty goal		
	score	proportion	Attempted	hit rate	score	proportion		Attempted	hit rate
1	2	2	2	1	3	3	3	3	4
Free throw percentage	Free Throws Attempted	free throw percentgetting older	sum of binary ball and three pointer	backboard	block shot	assist	Steal	fault	foul
4	4	3	2	2	2	1	1	5	3

Steps and methods for clustering in SPSS10.0 to open the database TABLE 1, click Analyze, Classify, Hierarchical Cluster input the number of clustering to 5 classes. According to principal component analysis in TABLE 2, of all the

statistical variables, main potential five major components, and five principal component represents the total variable of 83. 8% of the information, so make sure the statistical indicators of 5 classes.

The clustering results

Clustering results are shown in TABLE 3, the principal component 1 includes averaged binary shooting, assists, steals. Principal components including 2 binary ball scores, binary ratio, binary shots, the sum of binary and three-point shots, rebounds, blocks. Principal component 3 including goal, free-throw percentage, total score three shots, 3-point percentage, foul. Principal components including 4 penalty goal, the penalty goal proportion of total score, shots. Five principal components including errors. Principal component 1 contains the indexes and its correlation coefficient are shown in TABLE 3. Principal component 4 contains the indexes and its correlation coefficient are shown in TABLE 4. Principal component 5 ingredients - single index is the only mistake. The R composition clustering tree diagram, detailed instructions when class spacing is not at the same time, can put the statistical indicators for different class number. When class spacing equal to zero, each indicator is independent category, class spacing is smaller, get together for the same kind of index similarity degree is higher. When spacing is equal to 5, were basic divided into 10 types. When the class spacing equal to 15, into five types, also is the result of principal component analysis.

TABLE 3 : Binary shooting factor variable Pearson correlation coefficient ($\alpha = 0.01$)

	The ball better	assist	Steal	points per game	R ²
The ball better	1	0.699411	0.451215	0.702975	0.395404
P	-	0.000143	0.026889	0.000128	-
assist	0.699411	1	0.357825	0.695781	0.36706
P	0.000143	-	0.086021	0.00016	-
Steal	0.451215	0.357825	1	0.296794	0.139925
P	0.026889	0.086021	-	0.159032	-
points per game	0.702975	0.695021	0.296794	1	0.355611
P	0.000128	0.00016	0.159032	-	-

TABLE 4 : Frequency of penalty factor variables Pearson correlation coefficient ($\alpha = 0.01$)

	penalty goal	Free throw percentage	FTAfree throw attempts	R ²
penalty goal	1	0.83401	0.882511	0.737623
P	-	4.12E-07	1.17E-08	-
Free throw percentage	0.83401	1	0.851854	0.71073
P	4.12E-07	-	1.29E-07	-
FTAfree throw attempts	0.882511	0.851854	1	0.752797
P	1.17E-08	1.29E-07	-	-

The principal component analysis

Followed by principal component analysis of the research object, it is well known that the higher the proportion of the variance of main components indicate that the discrete degree of ingredients in each of the team, the greater the namely teams on the composition, the more obvious, the difference of the results to the team, of course, also the influence of the stronger. Therefore, from the perspective of the size of the main ingredients, all the teams to participate in the NBA playoffs are should attach great importance to all the team to the training of two points in throwing the ball. Because it is the main key points score, is the main components of the key data. After principal component analysis of clustering analysis, there are a few data are row behind the main component, they respectively are free-throw scores composition, composition of shots and failures. As we all know, the east and west teams in the NBA playoffs are very comprehensive technical team. Their data is very has the research value, to reflect the current international highest level of the overall trend of the development of basketball game. But as a result of national team has its own advantages and characteristics of the use of principal component to guide the team training and competition must be combined with their own characteristics, special skill training and improve.

Binary ball scores

The two-point shot score composition occupies very important position in the clustering analysis, we from the above TABLE display data can be seen clearly, scoring two goals is the main component of the second part of a proportion of the specific data is up to 22% of the total variance. This data is critical because it reflects one of the team's players in high level

high level is the main game scoring. This data is also important to a team of players in the game ball shots, rebounds, blocks which can have significant correlation is very intuitive to see from the above TABLE. As is known to all, for any basketball team in the NBA is very high-level league there, any team of players for the team and the team's shots it is a very long and difficult process. The NBA league is a competition is very fierce competition, especially into the playoffs, teams compete is very big, the competition is unusually fierce. Which in this case, how to let players can adapt to the rhythm of the game, still maintain a high efficiency goals, very stable. Because chance is very big in the game, players are hard to maintain a very stable field goal percentage. Therefore, can be completely by adding shots to make up for the inadequacy of shooting, and only shot have a scoring chance. Of course, this is a dialectical relationship, create good shots, as long as the athletes to build confidence, decisive and bold shot release, do not rely too much on a single focal point, only, can distract opponents defense, achieve the goal of rival small-scale operations, and win the game.

Turnovers

Error component analysis and cluster analysis of a very important data analysis. For the error component analysis, all the teams in the NBA playoffs data statistics shows error data is very obvious, the analysis of the data we have to consider the composition of the look is not only a simple one. The change rule of data information or single point data appear in the overall statistics of risk, to put it in the whole data for analysis. In the playoffs a mistakes occurrences and real data, we select the timberwolves, for this team, we from the point of view of statistical information data. The team averaged error number, and the all of the teams in the league error is relatively large number of components, many data shows that the playoffs the heat in the control of the ball ability still belong to the top teams in the NBA. But we found that the defender in the control time and control of the ball on the processing technology is need to make the necessary adjustments. Few games in the playoffs we see is the main cause of the failure of the Miami last with defender wade's ball control the timing of the timing and passing a lot of time processing has a lot to do. It is often the processing error led to the opponent to beat the last. In the heat of the many times the last three minutes of the game are often the problems.

Free throws

NBA playoff teams stand out in a team is Miami, a shot and free throws from the team scoring ratio has obvious advantage. But in terms of shots behind the 1 ~ 8 teams, showed that the strength of the attack, strength, the speed of moving also can't let the passive defense, all the defense is only in a disadvantageous position. The passive situation would have to foul or subconsciously, when defense initiative, under the favorable situation is almost impossible to take the initiative to foul. It is comply with the general logic of basketball sports.

CONCLUSION

Through the above analysis of the NBA basketball game in each link, using principal component analysis and clustering analysis, and applies this method to the comprehensive evaluation of the NBA basketball game. Principal component analysis (pca) and cluster analysis method for the NBA basketball game is very extensive in application of comprehensive evaluation, the research for improving our country on the principal component analysis (pca) and cluster analysis method research has the very vital significance. This research mainly through to the 2013 season in the NBA basketball game in his first season in pre-season technical data analysis of each project. By using the statistical methods of principal components analysis for each project, the multiple statistical index was clustering scores for binary ball shot composition, binary composition, free-throw scores composition, composition of shots, and the error component in the game. At the same time, according to each component to the importance of the impact on the overall variables combined with the characteristics of the NBA teams put forward a few research, the game players dribbling, passing, and widen the vision control aspects of the basic training to reduce errors and other technical points have very important guiding role.

REFERENCES

- [1] Sun Jilong, Wu Shengbao; The objectivity of CBA basketball game comprehensive score analysis [J], Journal of anhui normal university (natural science edition), **3**, (2003)
- [2] Jin Chuanjiang, Cai Zhongjia, Lv Shengjun, wang, Wang Xinghua; Application of Q type cluster of CBA basketball teams tactics ability of comprehensive evaluation [J], Journal of Beijing sports university, **4**, (2003).
- [3] Long Bin, Xiong Xibei; Men's sprint 300 m run different strength of blood lactic acid, blood urea, experimental study of changes of urinary protein [J], Journal of Beijing sports university, **5**, (2003).
- [4] Cui Bateer, Hu Xiaoyan, Cui Lei; College students' physical health evaluation system and research and development of exercise prescription [J], Journal of Beijing sports university, **7**, (2004).
- [5] Li Ruiwen, Cai Zhongjian, Jiang Bole; Sports a variety of forecasting methods of comparative study [J], Journal of Beijing sports university, **12**, (2004).
- [6] Jin Jiasheng; Comparative study of three kinds of commonly used sports comprehensive evaluation method - to 2003-2004 NBA outstanding in active service center competitive level of the [J], Journal of Beijing sports university, **2**, (2005).
- [7] Zhong Ping, Liu Yong; CBA basketball player lineup of offensive and defensive ability of give attention to two or

morethings fuzzy decision [J], Journal of Beijing sports university, **6**, (2005).

- [8] Wang Baocheng, Kuang LuBin, Tan Zhenbin; The basic theory of basketball player physical ability training and content [J], Journal of capital institute of physical education, **3**, (2001).
- [9] Wang Taohua, Zhao Guanghui, Wang Xinheng; the student non statistics and analysis of [J] intelligence factorson physical education achievements, Journal of Guangzhou Sports University, **2**, (2000).
- [10] Xu Bin; study of [J] application and tendency of football goal of sports science and technology, Hubei, **1**, (2005).