

评价 NBA 球员能力和薪资关系的数学模型

摘要

根据贝利提供的理论,一名 NBA 球员所得到的薪酬和他们在场上得多少分有直接关系,不过这并不是唯一需要考虑的因素,因为这和是否能够帮助球队带来胜利、在场上的投篮出手数、失误数、篮板数以及犯规数也有着关联。

为了探讨球员能力和薪资的关系,本文主要运用多元分析中的主成分分析法和因子分析法,借助于 SPSS 软件对数据进行分析,首先从衡量球员技术水平的得分、助攻、投篮命中率等10项指标出发,得出各指标与公因子的表达式

$ZX_i = aF_1 + bF_2 + cF_3 + \varepsilon_i$ 。用因子分析法对2011-2012赛季的8个球队现役球员的

综合能力进行评估分析,得到球员能力的综合指标模型,算出每个球员的综合得分。最后,我们将球员所得薪资与个人能力进行二次非线性回归,用 MATLAB 软件

拟合出两者之间的函数关系 $f(x) = p1 * x^2 + p2 * x + p3$ 。将算得的应得值与实际

所得值进行比较分析,得出检测值之间的误差,然后给出相关的合理解释。

关键词: 因子分析法、综合能力指标、回归分析

一、问题提出

科比、斯塔德迈尔、德克-诺维茨基等球员是 NBA 联盟里闪耀的球星，他们能够拿到动辄几千万的年薪也是无可厚非的事情。不过近日南犹他大学的经济学家教授大卫-贝利在进行的一项统计中，得出科比、小斯、诺维茨基等人实际上属于高薪低能的表现。他们的所得和所起到的表现并不能成为正比。

大卫-贝利对球员是否高薪低能的事情早有研究，为此他有着一套缜密的计算公式，而早在 2006 年的时候，他还和同伴一起撰写过一本名为“胜利的工资”的著作，在这本书中，贝利阐释了计算一名球员是否高薪低能的公式，这个公式被称之为“胜利产值”。根据贝利提供的理论，一名 NBA 球员所得到的薪酬和他们在场上得多少分有直接关系，不过这并不是唯一需要考虑的因素，因为这和是否能够帮助球队带来胜利、在场上的投篮出手数、失误数、篮板数以及犯规数也有着关联。

综合种种数据，经过测算，才会得出最终的“胜利产值”，以决定这名球员是否是高薪低能。在近日贝利给美媒体 CNBC 的一封邮件中，他表示：“在球场上，一支球队若想获得胜利，他们对球权的控制能力是关键，他们需要将球权转化为得分。而在 2011-12 赛季中，NBA 总共要为 990 场常规赛的胜利支付 19 亿美元，这意味着每场常规赛胜利合 194.6 万美元，根据每场胜利的花费和一名球员的获胜场数以及他的工资，就可以计算出他的性价比是否合格。”而令人感到意外的是，在贝利计算出的十大性价比低下的榜单上，竟然出现了科比、小斯、诺维茨基等人的名字，而且科比还位列这份榜单的第二位。

请收集 NBA 球员的表现数据，建立数学模型，合理评价球员的能力和薪资关系。

二、问题分析

NBA 球员的薪资高低是受个人能力、上场时间、球场经验、巨星效应等很多因素影响的，其中球员个人能力是评判一个球员薪资的一个最重要的指标。球员能力是由得分、篮板、助攻、抢断、盖帽等因素所决定。本文针对这一情况，建立了评价 NBA 球员能力和薪资关系的数学模型。首先，我们采用因子分析法，求得球员能力的综合得分，然后根据球员实际所得薪资与球员能力进行回归分析，得到拟合后的函数关系。最后我们对一个球员是否高薪低能进行了评判，模型的合理性得到了验证。

三、模型假设

- 1.假设收集的 NBA 球员数据均真实可靠；
- 2.假设球员所得的薪资跟所在的球队的战绩无关；
- 3.假设 NBA 30 支球队支付给球员的薪资不会超过联盟规定的工资帽。

四、符号说明及概念引入

- W_i : 第*i*个球员实际应得的薪资;
- A_i : 第*i*个球员的个人能力;
- ZXi : 第*i*个指标的个能力贡献;
- F_i : 第*i*个公因子;
- ε_i : 第*i*个外因子;
- λ_i : 第*i*个因子的方差值

五、模型的建立与求解

5.1 基于球员综合能力的模型：模块 I

针对球员能力与得分、篮板、助攻、盖帽、抢断、失误等十项个人数据，我们采用因子分析法对其进行分析。考虑到 *NBA* 球队众多，且每支球队人员组成差别不大，故我们选取题中提到的十名球员所在的8支球队进行分析，下面以篮网队为例，求出球员综合能力指标。在 *SPSS* 中进行因子分析的步骤如下：

为了确定本题中采用的因子分析适用性，我们采用了KMO和球形Bartlett检验。KMO检验球员指标间的偏相关是否较小，Bartlett球形检验是判断相关矩阵是否是单位阵，参见图5-1。

图 5-1

KMO 和 Bartlett 的检验		
取样足够度的 Kaiser-Meyer-Olkin 度量。		.797
Bartlett 的球形度检验 近似卡方		266.476
df		45
Sig.		.000

由Bartlett检验可以看出，球员指标间具有较强的相关性，而且KMO的统计量为0.797，大于0.7，说明各指标间信息的重叠程度还是比较高的。

由图 5-2 所示的公因子方差可知：提取的各个公因子几乎都在 80%以上，因此提取出的这几个公因子对各变量的解释能力是较强的。即提取的各指标对球员综合能力的评价程度较高。

图 5-2

公因子方差		
	初始	提取
得分	1.000	.943
篮板	1.000	.978
助攻	1.000	.911
抢断	1.000	.883
盖帽	1.000	.937
投篮命中率	1.000	.816
罚球命中率	1.000	.795
失误数	1.000	.953
出场次数	1.000	.763
出场时间（分钟）	1.000	.969

提取方法：主成份分析。

由下图可知，输出的结果，只有前三个特征根大于 1，前三个因子的方差贡献率为89.481%，因此选前三个因子已足够描述球员的综合能力水平。

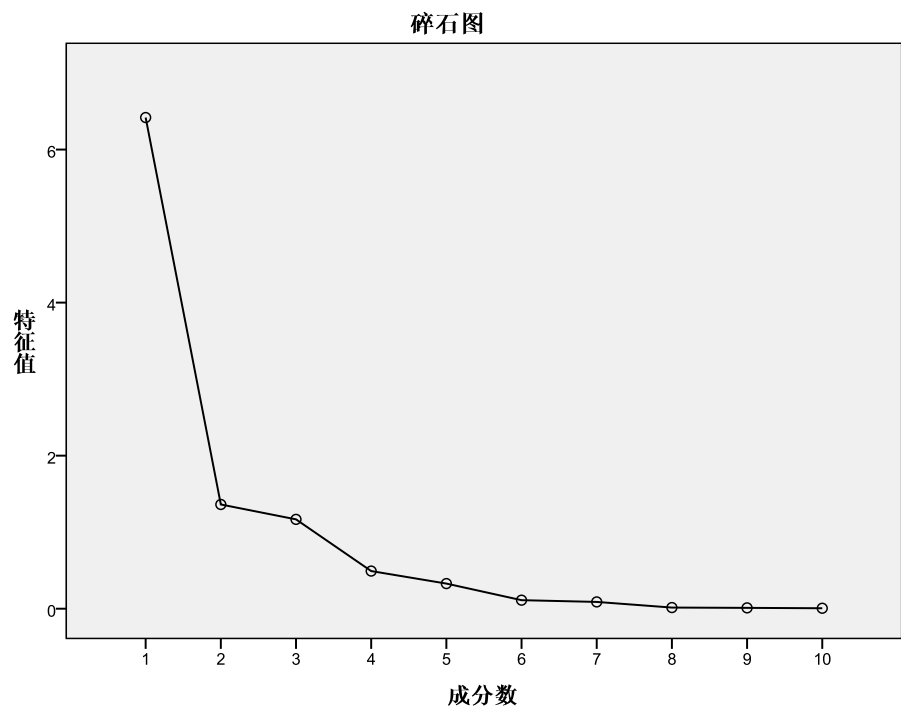
图 5-3

解释的总方差									
成份	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差的 %	累积 %	合计	方差的 %	累积 %	合计	方差的 %	累积 %
1	6.419	64.186	64.186	6.419	64.186	64.186	4.171	41.705	41.705
2	1.362	13.620	77.806	1.362	13.620	77.806	3.177	31.768	73.474
3	1.167	11.675	89.481	1.167	11.675	89.481	1.601	16.007	89.481
4	.492	4.918	94.399						
5	.328	3.283	97.682						
6	.111	1.115	98.796						
7	.089	.887	99.683						
8	.015	.148	99.831						
9	.011	.109	99.940						
10	.006	.060	100.000						

提取方法：主成份分析。

碎石图也进一步显示各因子的重要程度，参见图5-4。可见前三个因子的散点位于陡坡之上，而后七个因子散点成了平台，且特征根均小于1，因此至多考虑前三个公因子即可。

图 5-4



如下图 5-5，表示各因子对各球员指标变量的影响度。

图 5-5

成份矩阵^a

	成份		
	1	2	3
出场时间（分钟）	.979	-.058	-.083
得分	.947	-.161	.143
抢断	.927	-.153	-.012
失误数	.925	-.241	.197
出场次数	.867	-.009	-.105
篮板	.817	.252	-.497
盖帽	.765	.312	-.505
助攻	.732	-.456	.408
投篮命中率	.298	.795	.309
罚球命中率	.466	.501	.572

提取方法 :主成份。

a. 已提取了 3 个成份。

球员各个指标能力模型如下：

$$ZX_1 = 0.947F_1 - 0.161F_2 + 0.143F_3 + \varepsilon_1$$

$$ZX_2 = 0.817F_1 + 0.252F_2 - 0.497F_3 + \varepsilon_2$$

$$ZX_3 = 0.732F_1 - 0.456F_2 + 0.408F_3 + \varepsilon_3$$

$$ZX_4 = 0.927F_1 - 0.153F_2 - 0.012F_3 + \varepsilon_4$$

$$ZX_5 = 0.765F_1 + 0.312F_2 - 0.505F_3 + \varepsilon_5$$

$$ZX_6 = 0.298F_1 + 0.795F_2 + 0.309F_3 + \varepsilon_6$$

$$ZX_7 = 0.466F_1 - 0.501F_2 + 0.572F_3 + \varepsilon_7$$

$$ZX_8 = 0.925F_1 - 0.241F_2 + 0.197F_3 + \varepsilon_8$$

$$ZX_9 = 0.867F_1 - 0.009F_2 - 0.105F_3 + \varepsilon_9$$

$$ZX_{10} = 0.979F_1 - 0.058F_2 - 0.083F_3 + \varepsilon_{10}$$

在表达式中各指标标准化后的变量。**错误!未找到引用源。**表示特殊因子，是除了这3个公因子之外影响该变量的其他因素。原来设计了10个指标来表示球员综合能力水平，但是经过因子分析后，只需要三个因子即可描述影响球员综合能力水平状况。

本例中采用的是方差最大正交旋转法进行因子旋转，进行方差最大旋转后，旋转后的因子载荷矩阵如图 5-6 所示

图 5-6

旋转成份矩阵^a

	成份		
	1	2	3
助攻	.953	.011	.043
失误数	.899	.351	.144
得分	.852	.429	.182
抢断	.769	.531	.098
出场时间（分钟）	.730	.644	.146
出场次数	.611	.607	.146
篮板	.278	.941	.122
盖帽	.204	.934	.153
投篮命中率	-.061	.213	.876

罚球命中率	.325	.033	.829
-------	------	------	------

提取方法 :主成份。

旋转法 :具有 Kaiser 标准化的正交旋转法。

a. 旋转在 5 次迭代后收敛。

由图可以看出，第一公因子在 X_1 、 X_3 、 X_4 、 X_8 、 X_9 、 X_{10} 有较大的载荷，主要从得分、助攻、抢断、失误、出场次数和出场时间反映球员进攻能力，可以命名为进攻因子。第二公因子在 X_2 、 X_5 错误!未找到引用源。错误!未找到引用源。上有较大载荷，从篮板和盖帽方面反映球员防守能力，因此命名为防守因子。第三公因子在 X_6 、 X_7 上有较大载荷，表现为投篮命中率和罚球命中率，因此命名为稳定因子。与实际情况大致符合，各公因子的意义比较合理。

因子得分：公因子的得分系数函数不能通过矩阵变换的方法由因子载荷阵得到，只能采用估计的方法求得，本例采用的是回归法，将公因子表达为各变量的线性形式。因子得分系数矩阵如图 5-7 所示

图 5-7

成份得分系数矩阵

	成份		
	1	2	3
得分	.222	-.035	.010
篮板	-.176	.444	-.059
助攻	.403	-.288	-.043
抢断	.162	.061	-.060
盖帽	-.207	.458	-.030
投篮命中率	-.159	.023	.622
罚球命中率	.066	-.192	.582
失误数	.269	-.089	-.011
出场次数	.068	.146	-.022
出场时间（分钟）	.107	.132	-.037

提取方法 :主成份。

旋转法 :具有 Kaiser 标准化的正交旋转法。

构成得分。

可以直接写出各公因子的得分模型：

$$F_1 = 0.222ZX_1 - 0.176ZX_2 + 0.403ZX_3 + 0.162ZX_4 - 0.207ZX_5 - 0.159ZX_6 + 0.066ZX_7 + 0.266ZX_8 + 0.068ZX_9 + 0.107ZX_{10}$$

$$F_2 = -0.035ZX_1 + 0.444ZX_2 - 0.288ZX_3 + 0.061ZX_4 + 0.458ZX_5 + 0.023ZX_6 - 0.192ZX_7 - 0.089ZX_8 + 0.146ZX_9 + 0.132ZX_{10}$$

$$F_3 = 0.010ZX_1 - 0.059ZX_2 - 0.043ZX_3 - 0.060ZX_4 - 0.030ZX_5 + 0.622ZX_6 + 0.582ZX_7 - 0.011ZX_8 - 0.022ZX_9 - 0.037ZX_{10}$$

SPSS已经给出三个公因子的得分，保存在fac_1~fac_3中，按各因子对应的方差贡献率为权数计算如下综合统计量：

$$F = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} F_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} F_2 + \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} F_3$$

$$= 0.718F_1 + 0.152F_2 + 0.130F_3$$

在SPSS中用程序计算综合因子得分模型：

篮网

$$Comp\ score = 0.718 * fac1_1 + 0.152 * fac2_1 + 0.130 * fac3_1$$

根据以上原理，类似的，我们可求得以下七支球队的综合因子得分模型：

小牛

$$Comp\ score = 0.625 * fac1_2 + 0.255 * fac2_2 + 0.120 * fac3_2$$

奇才

$$Comp\ score = 0.705 * fac1_3 + 0.177 * fac2_3 + 0.118 * fac3_3$$

湖人

$$Comp\ score = 0.710 * fac1_4 + 0.173 * fac2_4 + 0.117 * fac3_4$$

尼克斯

$$Comp\ score = 0.651 * fac1_5 + 0.231 * fac2_5 + 0.118 * fac3_5$$

山猫

$$Comp\ score = 0.738 * fac1_6 + 0.262 * fac2_6$$

黄蜂

$$Comp\ score = 0.738 * fac1_7 + 0.262 * fac2_7$$

由以上模型可分别算出各个球队每名的综合得分，

小牛	2011-2012	球员		综合得分	
		德克-诺维茨基		1.1223	
		杰森-特里		0.7747	

		肖恩-马里昂		0.9071	
		文斯-卡特		0.3941	
		罗德里格-布博瓦		0.1615	
		德隆特-韦斯特		0.0351	
		伊恩-马辛米		0.1652	
		布兰登-莱特		0.0776	
		拉玛尔-奥多姆		-0.2828	
		杰森-基德		0.2294	
		布兰登-海伍德		0.1402	
		多米尼克-琼斯		-0.8356	
		易建联		-1.0329	
		布莱恩-卡迪纳尔		-1.0773	
		肖恩-威廉姆斯		-0.7785	
尼克斯	2011-2012	球员			
		卡梅罗-安东尼		1.3854	
		阿玛雷-斯塔德迈尔		0.5483	
		泰森-钱德勒		0.7491	
		兰德里-菲尔德斯		0.5335	
		伊曼-山姆珀特		0.9029	
		林书豪		0.5811	
		史蒂夫-诺瓦克		-0.1094	
		J. R. -史密斯		-0.1369	
		托尼-道格拉斯		-0.1157	
		比尔-沃克		-0.3354	
		拜伦-戴维斯		-0.317	
		贾里德-杰弗里斯		-0.5309	
		约什-哈雷尔森		-0.8037	
		迈克-毕比		-0.5469	
		雷纳多-巴尔克曼		-1.0246	
		杰罗姆-乔丹		-0.7798	
篮网	2011-2012	球员			
		德隆-威廉姆斯		2.2141	
		克里斯-亨弗里斯		0.3811	
		杰拉德-华莱士		0.9451	
		安东尼-莫罗		0.6158	
		马肖恩-布鲁克斯		0.7717	
		乔丹-法玛尔		0.3922	
		杰拉德-格林		-0.0384	
		桑迪亚塔-盖恩斯		0.3331	
		谢尔顿-威廉姆斯		-0.024	

		约翰-佩特罗		-0.081	
		乔丹-威廉姆斯		-0.3571	
		德肖恩-史蒂文森		-0.1736	
		布鲁克-洛佩斯		-0.6172	
		阿尔蒙-约翰逊		-0.4665	
		达米昂-詹姆斯		-0.5571	
		基斯-博甘斯		-0.7086	
		拉里-欧文斯		-0.5663	
		安德烈-埃梅特		-0.6881	
		杰里-史密斯		-0.8337	
		丹尼斯-霍纳		-0.5414	
湖人	2011-2012	球员			
		科比-布莱恩特		1.427	
		保罗-加索尔		0.8194	
		安德鲁-拜纳姆		0.3591	
		雷蒙-塞申斯		0.9071	
		慈世平		0.5808	
		马特-巴恩斯		0.1795	
		史蒂夫-布雷克		0.2297	
		乔丹-希尔		-0.6797	
		特洛伊-墨菲		-0.395	
		安德鲁-高德洛克		-0.6449	
		约什-麦克罗伯茨		-0.5023	
		德文-伊班克斯		-0.6898	
		达柳斯-莫里斯		-0.873	
		克里斯蒂安-艾恩加		-0.7179	
奇才队	2011-2012	球员			
		约翰-沃尔		1.9156	
		乔丹-克劳福德		0.852	
		凯文-塞拉芬		0.4557	
		特雷沃-布克		0.4106	
		克里斯-辛历顿		0.363	
		简-维塞利		0.1808	
		Shelvin-马克		-0.0336	
		安德雷-布拉切		-0.2703	
		拉沙德-刘易斯		-0.1951	
		内内-希拉里奥		-0.4715	
		卡地亚-马丁		-0.6176	
		莫里斯-埃文斯		-0.5413	
		詹姆斯-辛历顿		-0.3717	

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。

如要下载或阅读全文，请访问：

<https://d.book118.com/115321311324011334>