Discussion on the Existence of "Hot Hand Effect" in Competition

Solving practical problems using combinatorial mathematicstitle as needed

Zhang Junhua [1]

[1] *XiDian University, Xi'an, China, 710126*

Email: *ZhangJunhua@stu.xidian.edu.cn*

Abstract: "Hothand effect" is a topic often mentioned in the game. Take basketball as an example. If an athlete hits multiple times in a match, the player and the audience are more likely to believe that his next shot is more likely. Hit. This is the hot hand effect in the game. This paper analyzes and discusses the existence of the "hot hand effect" in the game by combining the relevant principles of mathematics. The recursive formula of combinatorial mathematics is used to explain the possible causes of the "hot hand effect". The analysis results were preliminarily verified by computer simulation experiments.

Keywords: hot hand effect; combinatorial mathematics; permutation and combination; recursive formula

比赛中「热手效应」存在性的探讨

使用组合数学原理解决实际问题

张俊华 [1]

[1] 西安电子科技大学，陕西 西安，中国，710126

Email: ZhangJunhua@stu.xidian.edu.cn

摘　要：「热手效应」是一个在比赛中被经常提及的话题，以篮球运动为例，如果一个运动员在比赛中连续多次命中，那么队员和观众更相信他的下一次出手更有可能命中。这就是比赛中的热手效应。本文通过组合数学的相关原理，对比赛中「热手效应」的存在性进行了分析和探讨，使用组合数学的递推公式，对「热手效应」可能出现的原因进行了说明。并通过计算机模拟实验的方法对分析结果进行了初步验证。

关键词：热手效应；组合数学；排列与组合；递推公式

1 引言

热手效应来源于篮球运动中。指比赛时如果某队员连续命中，其他队员一般会相信他“手感好”，下次进攻时还会选择他来投篮。无论是业余篮球爱好者还是职业球员，相信“手感”的人不在少数。他们用“手风顺”“手热”“着火”来形容连续命中投篮那种令人振奋的势头。但是，过去连续投进或投丢的球，真的会影响下一次出手会不会命中吗？换言之，所谓的“热手效应”（hot hand effect）是否存在？

1985 年，康奈尔大学的托马斯 •季洛维奇（Thomas Gilovich）、斯坦福大学的罗伯特 •瓦隆（Robert Vallone）和阿莫斯-特沃斯基（Amos Tversky）三位研究者对热手效应进行了研究。季洛维奇等人拿 NBA1980-1981 赛季费城 76 人队的真实数据进行了分析[1]，分别计算了球员们在投丢 1、2、3 次球之后下次投篮命中的概率，以及投中 1、2、3 次球之后下此投篮命中的概率。此外，他们还从康奈尔校队找来篮球选手进行受控的投篮实验，并做同样的分析。研究者指出，如果“热手效应”存在，那么连续命中 3 球后下一球的命中率应该比连续投丢 3 球后下一球的命中率高。然而经过计算，他们发现即便是这极端两种情况下，第四球的平均命中率都没有显著差异。

但是，既然有了统计和概率上的证明投篮的命中率与前几次投篮是否命中无关，那么为什么观众和球场上的球员仍然会认为某位球员当前的「手感好」呢？

2 使用递推公式对投篮命中结果分析

2.1 对问题的抽象

我们假设，在一场比赛中，一名队员投了 n 个球，则有 2^n 种可能的结果，每一次的投球结果或是投中，或是没有投中。将第 i 次投球投中记为 X_i ，将第 i 次投球未投中记为 O_i 。

设 B_n 表示 n 次投球，不包含连续投中的 X 和 O 的序列的子集，（即序列中不存在连续的两个 X ）。

我们把第 i 次投中记为 c_i ，并设 $C_n=\{c_1,c_2,\ldots,c_n\}$，因此，我们要解决的问题就是，不包含 $c_i,c_{i+1},i=1,2,\ldots,n-1$ 的 C_n 的集合有多少。

2.2 问题求解

设集合 B 中包含 c_n ，那么 $B-\{c_n\}$ 是 C_{n-2} 的一个子集，因为 c_{n-1} 一定不包含在 B 中。

如果集合 B 中不包含 c_n ，那么 $B-\{c_n\}=B$ 是 C_{n-1} 的一个子集，因此，对于 $n\geq 2$ ，有：

$$b_n=b_{n-1}+b_{n-2} \qquad (1)$$

特别的，当 $n=1$ 时，一次投球或投中或未投中，因此 $b_1=2$ 。当 $n=2$ 时，两次投球，要么没有投中的，或者第一次投中而第二次未投中，或者第二次投中而第一次未投中。因此 $b_2=3.$

设数列 $\{b_n\}$ 的母函数为数列 $G(x)$ 则：

$$\begin{aligned}
G(x)&=b_0+b_1x+b_2x^2+\cdots+b_nx^n+\cdots\\
&=1+x+(b_0+b_1)x^2+\cdots+\\
&\quad(b_{n-2}+b_{n-1})x^n+\cdots\\
&=1+x+b_0x^2+b_1x^2+b_2x^3+b_1x^3\cdots+x^n+\cdots\\
&\quad+b_{n-2}x^n+b_{n-1}x^n+\cdots\\
&=1+(x+b_1x^2+b_2x^3+\cdots+b_{n-1}x^n+\cdots)\\
&\quad+(b_0x^2+b_1x^3\cdots+b_{n-2}x^n+\cdots)\\
&=1+x(1+b_1x+b_2x^2+\cdots+b_{n-1}x^{n-1}+\cdots)\\
&\quad+x^2(b_0+b_1x+b_2x^2+\cdots+b_{n-1}x^{n-1}+\cdots)\\
&=1+(x+x^2)G(x)
\end{aligned}$$

解得：

$$G(x)=\frac{1}{1-x-x^2} \qquad (2)$$

利用待定系数法解得：

$$\begin{aligned}
G(x)&=\frac{\frac{1}{\sqrt5}}{x+\frac{1+\sqrt5}{2}}-\frac{\frac{1}{\sqrt5}}{x+\frac{1-\sqrt5}{2}}\\
&=\frac{1}{\sqrt5}\cdot\frac{1+\sqrt5}{2}\cdot\frac{1}{1-\frac{1+\sqrt5}{2}x}\\
&\quad-\frac{1}{\sqrt5}\cdot\frac{1-\sqrt5}{2}\\
&\quad\cdot\frac{1}{1-\frac{1-\sqrt5}{2}x} \qquad (3)
\end{aligned}$$

由幂级数展开公式 $\frac{1}{1-x}=1+x+x^2+\cdots+x^n+\cdots(-1<x<1)$ 得：

$$\begin{aligned}
G(x)=&\frac{1}{\sqrt5}\cdot\frac{1+\sqrt5}{2}\cdot\left[1+\frac{1+\sqrt5}{2}x+\left(\frac{1+\sqrt5}{2}\right)^2x^2+\cdots+\left(\frac{1+\sqrt5}{2}\right)^nx^n+\cdots\right]\\
&-\frac{1}{\sqrt5}\cdot\frac{1-\sqrt5}{2}\cdot\left[1+\frac{1-\sqrt5}{2}x+\left(\frac{1-\sqrt5}{2}\right)^2x^2+\cdots+\left(\frac{1-\sqrt5}{2}\right)^nx^n+\cdots\right]
\end{aligned} \qquad (4)$$

合并同类项，并与 $G(x)=b_0+b_1x+b_2x^2+\cdots+b_nx^n+\cdots$ 相比较，得：

$$b_n=\frac{1}{\sqrt5}\left[\left(\frac{\sqrt5+1}{2}\right)^{n+2}-\left(\frac{1-\sqrt5}{2}\right)^{n+2}\right] \qquad (5)$$

因此，如果尝试投球 10 次，在$2^{10}=1024$ 个可能的结果中，只有 $b_{10}=144$ 个结果不包含连续的投中。

2.3 结果分析

通过之前的计算结果可知，篮球比赛中球员不包含连续投中的比赛结果极少出现。在比赛中，连续命中（或连续投丢）相比于投中和投丢交替出现的情况更加让人印象深刻。研究者指出，人们可能对前者进行了记忆，而将投中投丢交替出现的情景不自觉地进行了忽略，这种 「选择性记忆」（selective amnesia）机制也有可能强化了人们对热手效应的认同。[2]

研究者们认为，人们之所以会“感知”到“热手效应”的存在，本质上对随机性的一种错觉。特沃斯基（Amos Tversky）和诺奖得主丹尼尔 •卡尼曼（Daniel Kahneman）曾总结：在实际生活中，人们会错误地将每次随机试

验之间独立的概率建立起联系。用掷硬币的例子来说，我们知道每次抛出得到正反面的概率都是 1/2，但总有人会认为如果连续几次都得到正面，那么下次得到反面的概率就会更大。这种以为在整体上符合期望的概率分布，在局部上也会符合相同的概率。这种将从大样本中得到的规律错误应用于小样本中的现象，被称为“小数定律”（law of small numbers）。季洛维奇等人曾在研究中招募学生球迷观看 21 投 11 中的投篮视频，其中一些视频中投篮进或不进被编辑成是随机的，但认为这些投篮是随机的人只有 32%。[1]

3 对真实比赛结果的分析

搜集整理 5 年内部分 NBA 球星的连续得分数量如下表所示：

表 1. 标准试验系统结果数据

连续进球数	汤普森	哈登	詹姆斯	欧文	杜兰特	安东尼
1	1073	1214	1115	943	950	918
2	476	485	578	411	431	361
3	195	202	288	170	203	140
4	79	81	150	79	91	74
5	35	34	65	34	48	22
6	14	16	28	7	18	9
7	13	2	21	10	12	2
8	5	5	2	4	11	4
9	3	0	3	2	1	2

以球星汤普森为例，对其得分情况进行分析：http://www.stat-nba.com/

从汤普森过去五年的赛场表现来看，其共投球 7301 次，投中 3343 次。假设没有热手效应这种情况，即命中的发生是随机的，而且每一次投球的命中概率相同，因此，3343 次投球和 3958 次失球的所有可能的顺序是均等的

$$P_0 = C(7301,3343) = \frac{7301!}{3343!3958!} \qquad (6)$$

将连续投中 r 次视作是 1 个排列单位 X^* 于是，我们计算 $3343-r$ 个 X，3958 个 O，一个 X^* 的排列数。

$$P_r = C(7301-r, 3343-r) * (7301-r) = \frac{(7301-r)!}{(3343-r)!\,3958!} * 7301 \qquad (7)$$

因此，在 7301 次投球中出现连续投中 r 球的期望次数为：

$$m = \frac{P_0}{P_r} \qquad (8)$$

利用 Python 程序对进球情况进行分析验证，得到结果：

至少连续进球次数	计算期望	实际连进数	
1	1530.45	1746	−0.12345
2	700.444	820	−0.1458
3	320.521	344	−0.06825
4	146.645	149	−0.01581
5	67.0828	70	−0.04167
6	30.6819	35	−0.12337
7	14.0308	21	−0.33187
8	6.41524	8	−0.1981
9	2.93273	3	−0.02242

3.2 真实情况与计算结果统计分析

看似不可能完成的 9 连进，通过排列统计，在 7301 次投球中，仍会发生大约 2.9 次。与汤普森的实际表现相吻合。**可见，所谓看似是奇迹的「出色手感」在大数据的统计下仍然为正常事件。**

可是，仔细观察计算的期望会发现，实际的连进次数，均比计算的进球期望稍高，将球员进球看做是进球结果的排列组合问题使得问题过于理想化，我们不能简单的将每一次的投球看成是完全随机互不干扰的事件，这五年内球员的能力成长，伤病，以及每场比赛的发挥情况，都会对真实的比赛结果造成影响。

4 计算机模拟验证

4.1 模拟进球结果

为了验证我们的预测结果是否正确，我们编写程序对投篮结果进行模拟预测，以上一例的球员汤普森的得分率 45.7% 为基础，使用程序模拟随机试验，记录每一次试验结果中的连进次数。

965	2885	1326	625	292	130	63	22	15	10	0	1	0	0	0	0	1	0	0	0	0	0
966	2975	1354	617	275	118	69	31	7	6	1	0	1	1	0	0	0	0	0	0	0	0
967	3120	1345	607	284	111	59	21	11	4	3	1	1	0	0	0	0	0	0	0	0	0
968	2995	1327	614	283	135	51	21	15	7	3	1	1	1	0	0	0	0	0	0	0	0
969	2949	1332	620	268	122	67	30	15	6	2	0	2	1	0	0	0	0	0	0	0	0
970	2908	1388	617	312	124	37	23	14	6	1	1	1	0	0	1	0	0	0	0	0	0
971	3042	1316	633	253	143	52	27	7	8	6	1	0	0	0	0	0	0	0	0	0	0
972	2909	1298	607	287	110	71	37	15	9	5	0	1	1	0	1	0	0	0	0	0	0
973	2894	1385	607	289	132	71	20	10	1	3	1	0	0	0	0	0	0	0	0	0	0
974	2997	1339	655	290	123	48	19	7	7	2	1	0	1	0	0	0	0	0	0	0	0
975	3015	1346	602	294	126	57	28	8	4	3	1	0	0	0	0	0	0	0	0	0	0
976	2908	1357	600	293	136	54	23	13	6	3	2	0	1	0	0	0	1	0	0	0	0
977	2890	1353	634	274	139	48	28	13	7	1	2	0	0	2	0	0	0	0	0	0	0
978	2897	1349	628	263	133	51	34	12	11	2	3	1	0	0	0	0	0	0	0	0	0
979	2976	1340	625	275	114	44	36	21	2	4	0	0	2	1	0	0	1	0	0	0	0
980	2939	1345	620	290	137	51	21	8	7	5	3	0	0	0	0	0	0	0	0	0	0
981	2964	1352	620	267	141	48	29	13	7	3	1	0	0	0	0	0	0	0	0	0	0
982	3008	1340	621	286	120	43	33	9	10	2	2	1	0	0	0	0	0	0	0	0	0
983	2950	1306	625	296	119	51	36	18	4	1	0	3	0	0	0	0	0	0	0	0	0
984	2934	1333	623	283	129	71	22	11	5	3	1	0	0	0	0	0	0	0	0	0	0
985	2938	1398	590	281	126	56	22	18	6	2	2	1	0	0	0	0	0	0	0	0	0
986	2939	1285	598	283	147	75	29	16	3	0	2	0	0	0	0	0	0	0	0	0	0
987	3013	1226	611	295	140	66	28	11	7	3	2	1	0	0	1	0	0	0	0	0	0
988	2998	1357	619	274	123	62	27	6	5	2	2	2	0	0	0	0	0	0	0	0	0
989	2981	1373	587	278	134	63	18	11	7	4	1	2	0	0	0	0	0	0	0	0	0
990	2904	1383	662	274	114	50	28	13	6	0	1	1	0	0	0	0	0	0	0	0	0
991	2926	1359	644	272	118	65	24	14	4	4	0	0	0	0	0	0	0	0	0	0	0
992	2920	1331	611	274	151	71	21	13	4	2	0	0	0	0	0	0	0	0	0	0	0
993	2974	1382	593	283	127	62	25	8	5	2	0	2	0	1	0	0	0	0	0	0	0
994	3006	1338	622	274	129	56	37	7	2	3	0	1	0	0	0	0	0	0	0	0	0
995	2859	1360	611	287	125	68	26	14	7	5	0	0	0	0	0	0	0	0	0	0	0
996	2965	1336	625	281	120	68	24	10	6	3	1	0	1	0	0	0	0	0	0	0	0
997	3045	1361	563	294	126	57	24	13	4	2	5	1	0	0	0	0	0	0	0	0	0
998	2823	1342	662	276	122	65	27	15	4	1	3	0	1	0	0	0	0	0	0	0	0
999	2936	1348	617	282	118	68	25	16	5	2	1	1	0	0	0	0	0	0	0	0	0
RESULT	2946.76	1349.723	614.982	280.932	128.713	58.676	26.756	12.312	5.682	2.625	1.138	0.538	0.265	0.112	0.055	0.028	0.006	0.005	0.001	0	0.001

图 1. 部分实验模拟结果

进行 1000 次模拟试验，每次模拟 7301 次投球（汤普森五年内投球数），部分模拟结果如图一。

模拟结果显示，按照如此概率进球，每 7301 次投球中连进 9 球的次数约为 2.625，与之前计算结果相符。在这 1000 次模拟中，几乎每次都会发生 10 连进，最高甚至达到了 20 连进。与汤普森在赛场上的表现无显著差异，可见，「热手效应」在真实球场上对球员的影响甚微。

5 总结

通过运用组合数学的递推公式，我们计算发现，诸如篮球之类的竞技比赛中，球员不包含连续投中的比赛结果极少出现。在比赛中，连续命中（或连续投丢）相比于投中和投丢交替出现的情况更加让人印象深刻。因此，所谓 「热手效应」可能只是人对随机现象的一种错误认知。

此外，通过对球员过去比赛的真实数据进行分析，以及通过计算机模拟验证，赛场上运动员的实际表现大致符合组合规律，「热手效应」对球员的比赛结果影响较小，进一步佐证了对热手效应存在性的分析。

5.2 不足之处

将每一次的投球结果，与当前赛场情况脱离开来讨论，问题的分析可能过于简化。

篮球比赛中一种常见的现象是，当球员命中一球后，对方便会加强防守，这有可能使本来存在的热手效应因防守等因素的变化而隐匿在最终数据背后。与此相比，排球则较少受这一因素的干扰：毕竟双方队员之间间隔了一张网，与篮球相比，较难进行针对某一球员的特殊防守。来自德国的研究者们采用职业排球联赛中的数据，采用与前述研究一样的方法，发现热手效应在排球中其实是存在的：当球员扣球成功后，他下一次扣球成功的概率会相应升高。

这又吸引到了篮球研究者的目光——如果通过更细致的数据将篮球比赛中的干扰因素也去除掉，是否能更好地验证投篮中“热手效应”是否存在？2014 年，来自哈佛大学的 3 名学分析了 NBA 2012-2013 赛季超过 8 万次投篮的数据。[4]在这一研究中，研究者对球赛剩余时间、两队分差等比赛变量、投篮距离、投篮方式等投篮变量、防守者与投篮者的距离及身高差等防守变量，以及不同运动员个体差异都进行细致控制，对投篮的条件概率进行预测。结果发现，投篮命中率可能的确受到一个细微但显著的“热手效应”影响。具体来说，在当前投篮之前的 4 次投篮中，每多命中 1 球，将大约使这次投篮的命中率提高 1.2%。

在新结果浮出水面的同时，对旧结果的反思也越来越多。2015 年，博科尼大学的约书亚 •米勒（Joshua B.Miller）等人提出，季洛维奇最初的研究中连续不中和连续命中后再一次投篮的“平均命中率”没有差异，反而正是“热手效应”存在的表现。[5] 他们主张，如果这一效应不存在，对命中率取平均值的做法将会引入偏倚——连续投篮不中后命中的平均命中率反而应当比连续投篮命中的平均命中率要高。但“热手效应”就此被“正名”了吗？还没有。想要断定它存在与否，仍需要更多的证据。

References 参考文献

[1] Gilovich, Thomas, Robert Vallone, and Amos Tversky."The hot hand in basketball: On the misperception of random sequences."Cognitive psychology 17.3(1985):295-314.

[2] Tversky, Amos, and Thomas Gilovich."The cold facts about the"hot hand"in basketball."
Anthology Statist. Sports 16(2005):169.

[3] Avugos, Simcha, et al."The "hot hand"reconsidered:A meta-analytic approach." Psychology of Sport and Exercise 14.1(2013):21-27.

[4] Bocskocsky, Andrew, John Ezekowitz, and Carolyn Stein."The hot hand:A new approach to an old "fallacy"."8th Annual Mit Sloan Sports Analytics Conference.2014.

[5] Miller, Joshua Benjamin, and Adam Sanjurjo."Surprised by the gambler's and hot hand fallacies? A truth in the law of small numbers."(2015).Jordan Ellenberg,"Hot Hands"in Basketball Are Real. Slate. com