商業分析期末報告

《三分天註定?》



報告大綱





研究動機與目的

Research motivation and purpose

研究動機與目的

近十年來三分球的出手命中數圖

	季後賽三分球總數	季後賽比賽總數		平均
2011-2012	2983	1	68	17.75595238
2012-2013	3502	1	70	20.6
2013-2014	3978	1	78	22.34831461
2014-2015	4139	1	62	25.54938272
2015-2016	4432	1	72	25.76744186
2016-2017	4597	1	58	29.09493671
2017-2018	4906	1	64	29.91463415
2018-2019	5402	1	64	32.93902439
2019-2020	6029	1	66	36.31927711
2020-2021	5831	1	70	34.3
2021-2122	5128	1	50	34.18666667



研究動機與目的

NBA冠軍賽正在進行中,我們注意到投三分球成為一種流行,蔚為風潮。於是我們想探討究竟三分球到底對勝負有沒有重大的影響呢?

我們將利用2020-2021賽季季後賽的數據,來探討三 分球時代帶來的影響是否會提高球隊的勝率。



Association rule learning

取2020-2021賽季季後賽出賽16支隊伍數據之平均三分出手、平 均三分命中、勝率及平均得分,並以平均三分出手、平均三分命 中及平均得分是否高於聯盟平均以及勝率是否高於五成四項條件, 篩選出符合的球隊及項目,並從各項目組合中分析四項數據之間 的關聯性。

2020-2021賽季季後賽16支隊伍各隊數據

Teams	G	3PA	3PM	WIN%	PTS
Portland	6	39.2	16.2	0.333	119.5
Utah	11	43.6	18	0.545	116.9
Philadelphia	12	28.4	11.2	0.583	116.3
Dallas	7	35	13.6	0.429	106
Denver	10	38.5	14.7	0.4	114.8
Brooklyn	12	35.6	13.6	0.583	112.5
LA	19	36.4	13.8	0.526	111.6
Pheonix	22	29.7	11.2	0.636	109
Boston	5	36.6	12.8	0.2	112.2
Atlanta	18	34.2	11.7	0.556	106.3
New York	5	30.4	10.4	0.2	97
Miami	4	35.8	12	0	98
Memphis	5	31.6	10.4	0.2	115
Milwaukee	23	35	11.2	0.696	110.3
Los Angeles	6	32.8	9.8	0.333	97.5
Washington	5	24.2	7	0.2	110
Average		34.3082353	12.48		110.324706

Teams	3PA	3PM	WIN%	PTS
Portland	1	1	0	1
Utah	1	1	1	1
Philadelphia	0	0	1	1
Dallas	1	1	0	0
Denver	1	1	0	1
Brooklyn	1	1	1	1
LA	1	1	1	1
Pheonix	0	0	1	1
Boston	1	1	0	1
Atlanta	0	0	1	0
New York	0	0	0	0
Miami	1	0	0	0
Memphis	0	0	0	1
Milwaukee	1	0	1	0
Los Angeles	0	0	0	0
Washington	0	0	0	0

各球隊四項數據,若該項數據高於平均則填1,低於平均則填0

teams	items
Portland	3PA,3PM,PTS
Utah	3PA,3PM,WIN%,PTS
Philadelphia	WIN%,PTS
Dallas	3PA,3PM,WIN%,
Denver	3PA,3PM,WIN%,
Brooklyn	3PA,3PM,WIN%,PTS
LA	3PA,3PM,WIN%,PTS
Pheonix	WIN%,PTS
Boston	3PA,3PM,
Atlanta	3PA,PTS
Miami	3PA
Memphis	PTS
Milwaukee	3PA,WIN%

右側欄:將上一頁四項數據中至少

有一項為1的球隊取出

左側欄:該球隊成功達到篩選條件

的品項

```
# -*- coding: utf-8 -*-
 Spyder Editor
 This is a temporary script file.
 from mlxtend.preprocessing import TransactionEncoder
 from mlxtend.frequent_patterns import apriori
 from mlxtend frequent patterns import association rules
 import pandas as pd
df_arr = [['3PA','3PM','PTS'],
['3PA','3PM','WIN%','PTS'],
                ['WIN%','PTS'],
['3PA','3PM'],
                ['3PA','3PM','PTS'],
                ['3PA','3PM','WIN%','PTS'],
['3PA','3PM','WIN%','PTS'],
                ['WIN%','PTS'],
                ['3PA','3PM','PTS'].
                ['WIN%'],
                ['3PA'],
                ['PTS'],
                ['3PA','WIN%']
 te = TransactionEncoder()
 df_tf = te.fit_transform(df_arr)
 df = pd.DataFrame(df_tf,columns=te.columns_)
frequent_itemsets = apriori(d'.min_support=0.25 use_colnames= True)
rules = association_rules(frequent_itemsetric = 'confidence'.min_threshold = 0.25)
rules = rules.drop(rules[rule.lift < 1].index)
rules.rename(columns = {'antecesents':'from','consequents':'to','support':'sup','confidence':'conf'},inplace = True)
rules = rules[['from','to','sup','conf','lift']]</pre>
 print(rules)
```

Python程式碼

度0.25、最小提升度1。

將項目欄中有包含項目的球隊取 出作為關連性分析樣本。 設定最小支援度為0.25、最小信賴

python程式結果

	from	to	aus	conf	lift	_
0	(3PM)	(3PA)	0.538462	1.000000	1.444444	1
1	(3PA)	(3PM)	0.538462	0.777778	1.444444	
6	(3PM)	(PTS)	0.461538	0.857143	1.238095	1
7	(PTS)	(3PM)	0.461538	0.666667	1.238095	
8	WIN%)	(PTS)	0.384615	0.714286	1.031746	1
9	(PTS)	(WIN%)	0.384615	0.555556	1.031746	
10	(3PM, 3PA)	(PTS)	0.461538	0.857143	1.238095	
11	(3PM, PTS)	(3PA)	0.461538	1.000000	1.444444	
12	(3PA, PTS)	(3PM)	0.461538	1.000000	1.857143	
13	(3PM)	(3PA, PTS)	0.461538	0.857143	1.857143	
14	(3PA)	(3PM, PTS)	0.461538	0.666667	1.444444	
15	(PTS)	(3PM, 3PA)	0.461538	0.666667	1.238095	

- 1.三分出手數高於平均與三分命中數高於平均具高關聯性
- 2.得分高於平均與三分命中數高於平均具高關聯性
- 3.得分高於平均與勝率高於五成具關 聯性
- **4.**當三分出手與三分命中皆高於平均時,得分有高機率高於平均

雖然由此分析無法直觀地看出提高三分出手數與勝率 超過五成是否具顯著關聯,但若單就進攻得分方面來 看,較高的三分出手數與命中數確實會對得分帶來正 面影響,因此若以拉高得分以增加獲勝機率為目標, 提升三分出手與命中數確實是有效的做法。

Paired Sample t test

取2020-2021賽季季後賽球隊逐場數據,分析季後賽

的勝方與敗方之間三分球進球數、命中率的平均數是

否有顯著差異。

場欠	勝方三分球運球數	· 敗 方三分球進球數	- 場交2	 勝方三分球進球 	· 敗方三分球運球	Ψ.
場次1		5	20 場次43		20	10
場次2		17	11 場次44		15	6
場次3		8	11 場次45		20	10
場次4		19	11 場次46		21	8
場次5		10	8場次47		13	14
場次6		9	7場次48		12	11
場次7		12	10 場次49		17	16
場次8		7	12 場次50		18	14
場次9		22	8場次51		6	8
場次10		12	16 場次52		20	11
場次11		17	12場次53		10	6
場次12		10	8場次54		12	14
場次13		18	13 場次55		19	19
場次14		9	2場次56		16	10
場次15		12	12場次57		8	14
場次16		19	8 場次58		12	11
場次17		13	9場次59		15	17
場次18		7	11 場次60		16	13
場次19		20	14 場次61		9	13
場次20		16	9 場次62		16	20
場次21		16	16 場次63		7	9
場次22		20	13 場次64		12	10
場次23		13	11 場次65		20	21
場次24		12	13 場次66		15	12
場次25		17	8場次67		7	12 9
場次26		19	13 場次68		13	20
場次27		15	9場次69		6	13
場次28		10	13 場次70		8	8
場次29		16	14 場次71		12	10
場次30		13	5 場次72		15	
場次31		9	12場次73		4	- 5
場次32		17	10場次74		15	0
場次33		15	11 場次万		10	9
場次34		20	21 場次76		13	9 5 9 9 8
場次35		14	12 場次77		17	12
場次36		9	8場次78		9	15
場次37		9	12場次79		17	12
場次38		17	14 場次80		11	
場次39		14	12 場次81		20	10
場次40		15	15 場次82		14	9
場次40場次41		18	10 場次83		7	16 9 9 7
場次42		10	10 場次84		14	13
*#6-15,42		10	場次85		6	6

樣本數:季後賽共85場



Ho: 勝方和敗方三分球平均進球 數皆相等

Ha: 勝方和敗方三分球平均進球 數不相等

分組方式:勝方、敗方各一組

 $\alpha:0.05$

t 檢定:成對母體平均數差異檢定		
	勝方三分球進球數	敗方三分球進球數
平均數	13.49411765	11.42352941
變異數	19.99103641	13.88991597
觀察值個數	85	85
皮耳森相關係數	0.236624116	
假設的均數差	0	
自由度	84	
t 統計	3.744196604	
P(T<=t) 單尾	0.00016526	
臨界值:單尾	1.663196679	
P(T<=t) 雙尾	0.000330519	
臨界值:雙尾	1.988609667	

勝方三分球平均進球數:13.49顆

敗方三分球平均進球數:11.89顆

T值=3.744

顯著性P值(雙尾)0.00033 < 0.05

故拒絕Ho,勝方之三分球平均 進球數顯著高於敗方。



場次 上 勝	方三分球命中率。	敗方三分球命中率*	場次二	勝方三分球命中之	敗方三分球命中率_
場次1	16.13%	40.00%	場次43	46.51%	27.78%
場次2	47.22%	27.50%	場次44	38.46%	20.00%
易-欠3	23.53%	36.67%	場次45	42.55%	34.48%
易-次4	47.50%	30.56%	場次46	50.00%	29.63%
易次5	31.25%	40.00%	場次47	38.24%	35.00%
場次6	32.14%	26.92%	場次48	46.15%	36.67%
易次7	35.29%	33.33%	場次49	34.00%	38.10%
易次8	35.00%	25.53%	場次50	47.37%	32.56%
易次9	41.51%	28.57%	場次51	19.35%	25.00%
易-欠10	42.86%	48,48%	場次52	51.28%	36,67%
易次11	44.74%	35,29%	場次53	47.62%	26,09%
易-次12	30,30%	30.77%	場次54	46.15%	34.15%
易-欠13	52.94%		場次55	52.78%	43.18%
易次14	42,86%		場次56	34.04%	30,30%
易-次15	36.36%	27.27%	場次57	36.36%	37.84%
易次16	48.72%		場次58	30.00%	40.74%
易-次17	36.11%		場次59	40.54%	40,489
易-次18	25.00%	37.93%	場次60	37.21%	40.639
易-次 19	52.63%	31.11%	場次61	34.62%	43,339
易-欠20	59.26%		場次62	40.00%	37.049
易-次21	41.03%	42.11%	場次63	21.21%	30,009
易-欠22	51.28%		場次64	41.38%	32,26%
易-欠23	30.95%	33.33%	場次65	51.28%	47.73%
易次24	36,36%		場次66	41.67%	30.779
易次25	51.52%		場次67	25.93%	32.149
易-次26	44.19%		場次68	40.63%	42,559
易次27	38,46%		場次69	23.08%	38.249
易-欠28	28,57%		場次70	25.00%	22,229
場-欠29	59,26%		場次71	35.29%	31.259
易-次30	39.39%		場次72	36.59%	25,009
場次31	37,50%		場次73	20.00%	16.139
易-次32	50.00%		場-次74	51.72%	
易次33	46.88%		場次75	33.33%	34.629
易次34	45,45%		場次76	34.21%	20.519
易次35	33,33%		場次77	54.84%	30,779
易次36	33,33%		場-次78	31.03%	39,479
易次37	26.47%		場次79	37.78%	37,509
易-次38	38,64%		場次80	32,35%	44,449
易次39	38.89%		場次81	50.00%	29.03%
易次40	44.12%		場次82	38,89%	29.03%
場次41	51.43%		場次83	24.14%	30.43%
易次42	29.41%		場次84	50.00%	
	23.41%	32.33 %	場次85	22,22%	24.00%

樣本數:季後賽共85場



Ho: 勝方和敗方三分球平均命中率皆相等

Ha: 勝方和敗方三分球平均命中 率不相等

分組方式:勝方、敗方各一組

 $\alpha:0.05$

t 檢定:成對母體平均數差異檢定		
	勝方三分球命中率	敗方三分球命中率
平均數	0.389838633	0.332166478
變異數	0.009599169	0.006637082
觀察值個數	85	85
皮耳森相關係數	0.09704609	
假設的均數差	0	
自由度	84	
t 統計	4.387414449	
P(T<=t) 單尾	1.65522E-05	
臨界值:單尾	1.663196679	
P(T<=t) 雙尾	3.31045E-05	
臨界值:雙尾	1.988609667	

勝方三分球平均命中率:38.98%

敗方三分球平均命中率:33.21%

T值=4.3874

顯著性P值(雙尾)0.00003 < 0.05

故拒絕Ho,勝方之三分球平均 命中率顯著高於敗方。



勝方的三分球進球數和命中率皆顯著高於敗方,且命中率的P值更低,勝敗兩方的差異更為明顯。

因此我認為**提高三分球進球數和命中率,皆有更大的機**

會獲勝,球隊則應優先提升命中率,重量更重值。

4.

線性&邏輯思迴歸

Linear&Logistic regression

線性&邏輯思迴歸

本分析2020-2021賽季季後賽球隊逐場數據,利用的

數據為:主、客場的得分、主客場三分球命中率、勝負、

三分球命中率、三分球出手,整理為表格如下:



線性&邏輯思迴歸

			√ J	* 110141	C_ICAIV	1_1/11/12											
4	Α	В	С	D	Е	F	G	Н	1	J	K	L	М	N	0	Р	
1 (JAME 🔻	GAME_	GAME_	HOME -	VISIT -	SEASON 🗷	TEAM_ID_	PTS_home 💌	FG3_PCT	TEAM_	PTS_away 🔽	FG3_PCT_	HOME_T_	三分命中(主-客)	主場三分命中率大於客場	主場三分出手大於客場	
078 2	2021/7/20	42000406	Final	1.61E+09	2E+09	2020	1610612749	105	0.222	1.61E+09	98	0.24	1	-0.018	0	1	
079 2	021/7/17	42000405	Final	1.61E+09	2E+09	2020	1610612756	119	0.684	1.61E+09	123	0.5	0	0.184	1	0	
080 2	2021/7/14	42000404	Final	1.61E+09	2E+09	2020	1610612749	109	0.241	1.61E+09	103	0.304	1	-0.063	0	1	
081 2	2021/7/11	42000403	Final	1.61E+09	2E+09	2020	1610612749	120	0.389	1.61E+09	100	0.29	1	0.099	1	1	
.082	2021/7/8	42000402	Final	1.61E+09	2E+09	2020	1610612756	118	0.5	1.61E+09	108	0.29	1	0.21	1	1	
083	2021/7/6	42000401	Final	1.61E+09	2E+09	2020	1610612756	118	0.324	1.61E+09	105	0.444	1	-0.12	0	0	
		42000306		1.61E+09	2E+09	2020	1610612737	107		1.61E+09	118	0.378	0		0	0	
085	2021/7/1	42000305	Final	1.61E+09		2020	1610612749	123		1.61E+09	112	0.395	1	0,000	0	0	
086 2	2021/6/30	42000316	Final	1.61E+09	2E+09	2020	1610612746	103	0.308	1.61E+09	130	0.548	0		0	1	
087 2	2021/6/29	42000304	Final	1.61E+09	2E+09	2020	1610612737	110	0.342	1.61E+09	88	0.205	1	0.137	1	0	
088 2	2021/6/28	42000315	Final	1.61E+09	2E+09	2020	1610612756	102	0.346	1.61E+09	116	0.333	0	0.013	1	0	
089 2	2021/6/27	42000303	Final	1.61E+09	2E+09	2020	1610612737	102	0.405	1.61E+09	113	0.379	0	0.026	1	1	
090 2	2021/6/26	42000314	Final	1.61E+09	2E+09	2020	1610612746	80	0.161	1.61E+09	84	0.2	0	-0.039	0	1	
091 2	2021/6/25	42000302	Final	1.61E+09	2E+09	2020	1610612749	125	0.366	1.61E+09	91	0.25	1	0.116	1	1	
092 2	2021/6/24	42000313	Final	1.61E+09	2E+09	2020	1610612746	106	0.353	1.61E+09	92	0.313	1	0.04	1	1	
093 2	2021/6/23	42000301	Final	1.61E+09	2E+09	2020	1610612749	113	0.222	1.61E+09	116	0.25	0	-0.028	0	1	
094 2	2021/6/22	42000312	Final	1.61E+09	2E+09	2020	1610612756	104	0.231	1.61E+09	103	0.382	1	-0.151	0	0	
095 2	2021/6/20	42000311	Final	1.61E+09	2E+09	2020	1610612756	120	0.406	1.61E+09	114	0.426	1	-0.02	0	0	
096 2	2021/6/20	42000207	Final	1.61E+09	2E+09	2020	1610612755	96	0.321	1.61E+09	103	0.259	0	0.062	1	1	
.097 2	2021/6/19	42000217	Final	1.61E+09	2E+09	2020	1610612751	111	0.308	1.61E+09	115	0.417	0		0	1	
_		42000206		1.61E+09	2E+09	2020	1610612737	99		1.61E+09	104	0.414	0		0	1	
_		42000226		1.61E+09	2E+09	2020	1610612746	131	0.513	1.61E+09	119	0.477	1		1	0	
100 2	2021/6/17	42000216	Final	1.61E+09	2E+09	2020	1610612749	104	0.212	1.61E+09	89	0.3	1	-0.088	0	1	
101 2	2021/6/16	42000205	Final	1.61E+09	2E+09	2020	1610612755	106	0.433	1.61E+09	109	0.346	0		1	1	
102 2		42000225 -21季後賽		1.61E+09 +	2E+09	2020	1610612762	111	0.37	1.61E+09	119	0.4	0	-0.03	0		L

線性迴歸

本場的得分與三分球命中率有無相關

採用數據

應變數:得分

自變數:三分球命中率



線性迴歸

係數a

		非標準	化係數	標準化係數		
模型		В	標準錯誤	β	Т	顯著性
1	(常數)	83.519	4.482		18.636	<.001
	FG3_PCT_home	80.750	12.076	.578	6.687	<.001

a. 應變數: PTS_home

係數a

		非標準化係數		標準化係數		
模型		В	標準錯誤	β	Т	顯著性
1	(常數)	75.600	4.058		18.630	<.001
	FG3_PCT_away	90.598	10.817	.664	8.375	<.001

a. 應變數: PTS_away

上:主場 下:客場

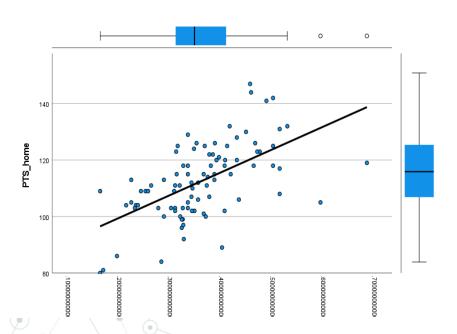
兩個分析結果顯示:

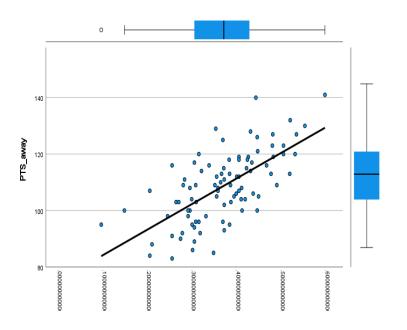
- 1.顯著性都小於0.001 表示其具有分析的 能力
- 2.標準化係數為正,得 分與三分球命中率 為正相關

線性迴歸

主場:





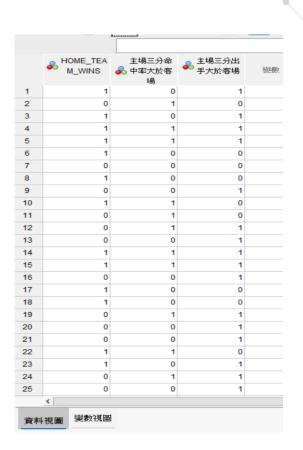


以勝利為應變數,主場 三分球命中率和三分出 手是否大於客場為共變 數做邏輯思迴歸預測

先將資料整理成二元資料,

1為主場勝利、主場大於,0為客場勝利、主場不大於,並利用SPSS

內建的二元邏輯思迴歸分析 做預測



模型係數的 Omnibus 檢定

		卡方檢定	自由度	顯著性	
步驟1	步驟	8.123	2	.017	
	區塊	8.123	2	.017	
	模型	8.123	2	.017	

分類表a

				HOME TE			
		觀察值		HOME_TE	4M_VVIIV5 1	正確百分比	
ĺ	步驟1	HOME_TEAM_WINS	0	26	12	68.4	
			1	21	32	60.4	
		整體百分比				63.7	

可以看到此模型顯著性 小於0.05,達到分析預 測水準,且其預測正確 率為63.7%

根據分析結果三分球出手數與當場是否勝利在此分析並無法達到分析水準(顯著性大於0.05),但就命中率而言對於勝利的預測是正確且有達到分析水準的,可以發現三分球命中率較小的那隊,其勝率只有30.6%。

								EXP(B) 的 95% 信報區間	
		В	S.E.	Wald	自由度	顯著性	Exp(B)	TR	上限
步驟 1 *	主場三分命中率大於客場(1)	-1.185	.450	6.946	1	.008	.306	.127	.738
	主場三分出手大於客場(1)	.358	.449	.637	1	.425	1.431	.594	3.448
	常數	.817	.392	4.345	1	.037	2.264		

少粮1上輸入的變數:[%1:,1:

依照邏輯思迴歸的分析結果,我們可以得知三分球命中率若超過敵隊可以做為左右勝負的關鍵,但是三分球的出手數超過敵隊則可能與本場勝負較無關聯。



Conclusion



結論

透過關聯性分析,高得分和三分球的出手數與進球數 有高度相關,相對樣本T檢定和迴歸分析皆顯示三分球 進球數和命中率皆有助於提高勝率,出手數雖對得分 有影響,勝場的影響卻不明顯。

因此,我們認為**三分球是左右球隊獲勝的關鍵**,三分 球時代除了流行且有實質上的效益,值得效法。



References



參考資料

https://www.kaggle.com/datasets/nathanlauga/nba-games?select=games.csv

https://www.kaggle.com/code/xuannaselli/it-s-raining-threes-in-nba-is-it-worth-it

https://www.nba.com/stats/teams/traditional/?sort=GP&dir=-1&Season=2020-21&SeasonType=Playoffs

https://www.google.com/search?q=20-21+nba+playoffs&oq=20&aqs=chrome.2.69i59l3j69i57j69i59j69i61l3.2164j0j7&sourceid =chrome&ie=UTF-8#sie=lg;/g/11jn2 x2b4;3;/m/05jvx;mt;fp;1