

2018-2021の日本プロ野球試合データから見た ホームアドバンテージの要因

東京理科大学 経営学部 朝日研究室

圓谷 誠

目次

1. 研究背景
 2. 研究内容について
 3. 使用データ概要
 4. 分析手法
 5. 基礎集計
 6. 基礎分析
 7. 分析
 1. 仮説Ⅰ 観衆因子（ロジスティックス回帰分析、相関分析）
 2. 仮説Ⅱ ルール因子（ウィルコクソンの符号順位和検定）
 3. 仮説Ⅲ 習熟度因子（t 検定）
 8. 考察
 9. まとめ
 10. 今後の課題
- Appendix
参考文献

1. 研究背景

- ◆ 日本プロ野球に所属する球団は勝利と集客を求めている
 - 年143試合の公式戦を通じたペナントレース（リーグ戦）を2リーグ制（セ・リーグ、パ・リーグ）で行っており、リーグ戦優勝を目指している
 - 主な収入源としてチケット販売があり、観客動員数は球団にとって非常に重要である

その中で、基礎分析から以下のことが分かっている

- ◆ 勝率の高さが、集客に影響を与えている
 - ◆ 優勝チームほどホーム勝率が高い傾向
- ⇒ホームアドバンテージを活かすことが球団にとって重要

先行研究では、ホームアドバンテージの有意性は認められているもののその要因について決定的な手掛かりは得られていない

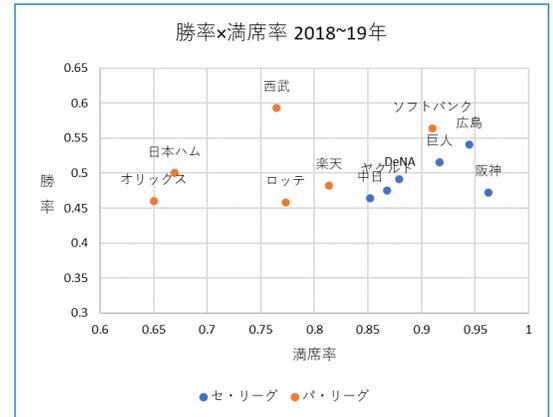


図1 2018~19年 勝率×満席率 散布図

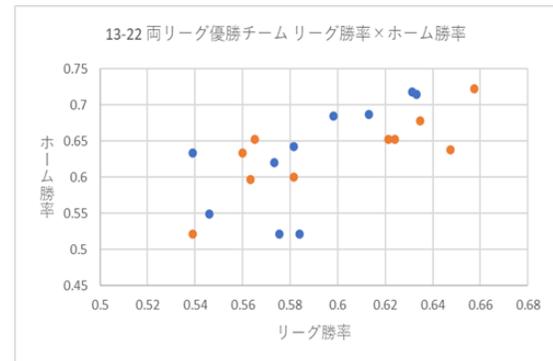


図2 過去10年両リーグ優勝チーム 勝率 散布図

2. 研究内容について

研究目的：ホームアドバンテージの要因を明らかにすることで、

- ① ホーム勝率を伸ばす
- ② アウェイ環境との違いを明確にすることで、戦い方に工夫をもたらす
- ③ ①,②により勝率向上させ、観客動員数・チケット販売収入を向上させる

仮説

- I. **観衆因子**がホームアドバンテージに影響を与えている
⇒球場内の観衆の数やそこから生まれる熱量が影響
- II. **ルール因子**がホームアドバンテージに影響を与えている
⇒ルール上、ホームチームが後攻になることが影響
- III. **習熟度因子**がホームアドバンテージに影響を与えている
⇒シーズンの中の約半分がホーム球場で行われることが影響

-先行研究との違い-

- 2018年以降NPBのデータを用いたホームアドバンテージの研究が少ない⇒研究の更新
- コロナ禍の入場規制により観衆因子について、より詳細な研究ができる
- 習熟度因子について、勝敗だけでなく攻守の指標を用いて分析

3. 使用データ概要

データスタジアム株式会社が提供する。
過去4年分（2018-2021）の日本プロ野球のデータ

日本プロ野球（NPB） 過去4年分の試合データ（2018-2021）

- 試合結果や打席投球情報など
- 人の目視によるデータ

表 1 チーム成績データ例

日付	ホームチーム名	アウェイチーム名	球場名	試合種別	試合種別詳細	ホームチーム得点	アウェイチーム得点	試合終了イニング	試合終了表
2020/11/14	D e N A	巨人	横浜	公式戦	セ・リーグ公式戦	5	4	9裏	

表 2 投球情報データ例

球速	投球位置座標X	投球位置座標Y	投球位置区域	投球結果	打撃結果	打球種別	打球位置座標X	打球位置座標Y	打球処理守備位置	打球強度
150	156	104	8	結果打球	凡打	ゴロ	133	198	二	弱

4. 分析手法

使用分析ツール：Excel、Alkano



図 3 Alkanoワークスペース

5. 基礎集計（入場者数について）

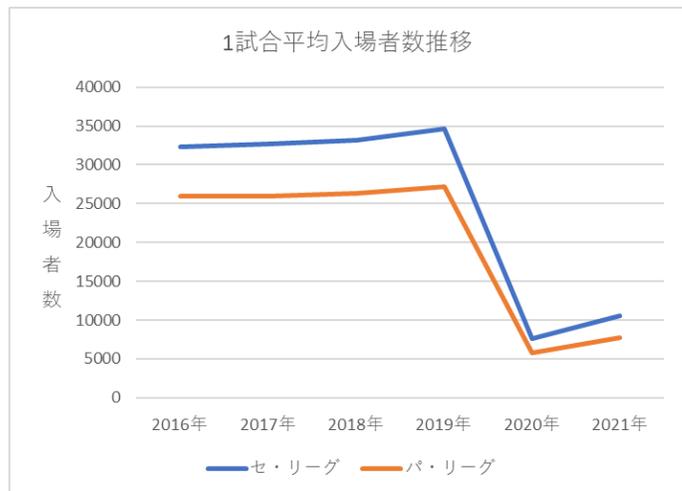


図4 NPB1試合平均入場者数推移

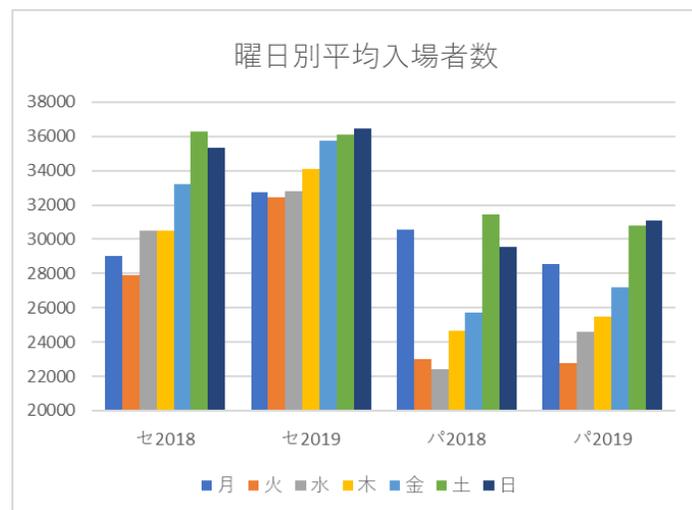


図5 NPB曜日別平均入場者数

2020年以降新型コロナウイルスの影響により入場制限などもあり入場者数は激減している

リーグごと、土日と平日などで入場者数には差が出る

リーグにより入場者数の差
また傾向にも違いがある
⇒入場者数に関わるものは
リーグ別で分析する必要性

5. 基礎集計（勝率について）

表3 セ・リーグ2018~21年 戦績

チーム	ホーム試合数	勝率	ホーム勝率	アウェイ勝率	ホームゲーム勝利数	敗北数	引分数
巨人	287	0.499216	0.50729927	0.49113356	139	135	13
阪神	286	0.513385	0.503623188	0.523147479	139	137	10
広島	286	0.497782	0.549450549	0.446112519	150	123	13
中日	285	0.457479	0.539033457	0.375924215	145	124	16
DeNA	287	0.477976	0.519855596	0.436095969	144	133	10
ヤクルト	285	0.544847	0.514492754	0.575200292	142	134	9
集計			0.52229247	0.474602339			

表4 パ・リーグ2018~21年 戦績

チーム	ホーム試合数	勝率	ホーム勝率	ホーム以外勝率	ホームゲーム勝利数	敗北数	引分数
ソフトバンク	286	0.531603	0.564575646	0.498631268	153	118	15
日本ハム	285	0.456194	0.5	0.412388695	134	134	17
西武	286	0.515929	0.578181818	0.453675325	159	116	11
楽天	286	0.496994	0.432727273	0.561260823	119	156	11
ロッテ	286	0.494913	0.48	0.509825382	132	143	11
オリックス	286	0.519669	0.524344569	0.514993492	140	127	19
集計			0.51330488	0.491795831			

全体の勝率に比べ、ホームゲーム勝率の方が高い傾向にある

5. 基礎集計（ホーム球場について）

球団別の使用球場の割合

表5 セ・リーグ18-21年使用球場割合（最大5%以下の球場は非表示）

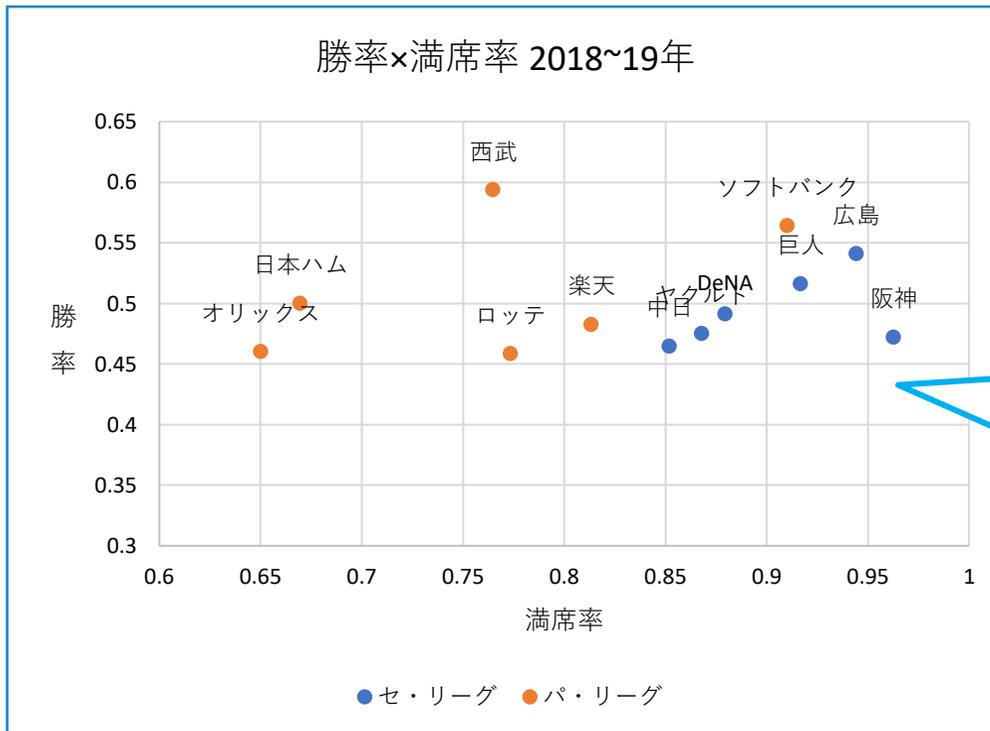
セ・リーグ	ナゴヤドーム	マツダスタジアム	横浜	京セラD大阪	甲子園	神宮	東京ドーム
DeNA	8.74%	9.11%	47.72%	2.19%	7.83%	9.29%	10.02%
ヤクルト	8.62%	8.62%	8.62%	2.87%	7.90%	46.50%	8.62%
巨人	8.84%	8.84%	8.30%	2.17%	9.39%	8.30%	45.49%
広島	8.93%	49.54%	8.38%	3.28%	6.74%	8.56%	9.47%
中日	49.18%	8.93%	8.74%	1.64%	7.83%	8.93%	8.20%
阪神	8.89%	9.07%	8.53%	5.81%	44.46%	8.53%	9.44%

表6 パ・リーグ18-21年使用球場割合（最大5%以下の球場は非表示）

パ・リーグ	ZOZOマリン	ほっと神戸	メットライフ	ヤフオクドーム	楽天生命パーク	京セラD大阪	札幌ドーム
オリックス	8.42%	5.73%	8.78%	8.78%	8.24%	44.27%	8.24%
ソフトバンク	9.11%	0.55%	8.74%	46.81%	8.93%	8.93%	7.29%
ロッテ	49.37%	0.90%	8.29%	8.47%	8.65%	8.83%	7.21%
楽天	9.44%	1.81%	7.99%	8.53%	48.46%	7.26%	7.99%
西武	9.11%	1.28%	47.91%	7.65%	8.93%	8.20%	6.56%
日本ハム	8.74%	1.09%	9.11%	8.20%	8.93%	7.47%	42.44%

年あたり40%程度の試合を行うホーム球場を各チーム持っている

6. 基礎分析



全体で見るとバラバラだが、
セ・リーグのみでの相関係数は**0.48**
パ・リーグのみでは**0.47**というように
正の相関が見られる

図1 2018~19年 勝率×満席率 散布図 (再掲)

入場者数に関わるものはリーグ別にした分析をする必要がある

6. 基礎分析

過去10年のリーグ戦優勝チームの戦績を基に
リーグ勝率（全試合）とホーム勝率（ホーム主催試合）の関係を分析

なお相関係数は
両リーグで0.68725359
セ・リーグ0.672631355、
パ・リーグ0.777434215
であり高い正の相関関係にあることが分かる



優勝チームは特にホーム勝率が高く
ホームアドバンテージを活かしており、
その点がリーグ優勝に繋がっていると考えられる

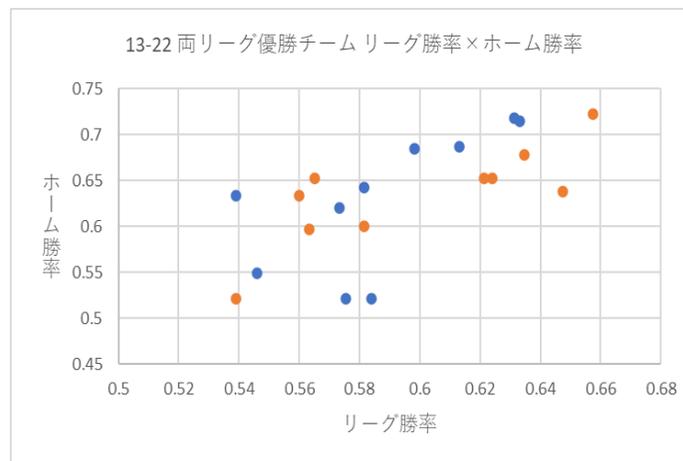


図2 過去10年両リーグ優勝チーム勝率 散布図
(再掲)

6. 基礎分析

ホームアウェイ勝率の差の検定

⇒今回扱うデータにホームアドバンテージはあるのか分析

使用データ：2018-2021の試合データ

表7 使用データ例

ホーム戦績	勝ち	引き分け	負け	合計	ホーム勝率	ビジター戦績	勝ち	引き分け	負け	合計	ビジター勝率
広島	51	1	20	72	0.718309859		38	1	32	71	0.542857143

帰無仮説：ホーム勝率=アウェイ勝率 対立仮説：ホーム勝率>アウェイ勝率
 という仮説を基に t 検定を実施（対応している2標本による平均の検定）

セ・リーグ2018-2021		
	ホーム勝率	アウェイ勝率
平均	0.53318568	0.460244874
分散	0.006717621	0.005545866
観測数	24	24
ピアソン相関	0.063809389	
仮説平均との差異	0	
自由度	23	
t	3.334412587	
P(T<=t) 片側	0.001440536	
t 境界値 片側	1.713871528	
P(T<=t) 両側	0.002881071	
t 境界値 両側	2.06865761	

図6 セ・リーグ勝率の差の検定

パ・リーグ2018-2021		
	ホーム勝率	アウェイ勝率
平均	0.52918004	0.47763149
分散	0.007689336	0.0048857
観測数	24	24
ピアソン相関	0.229828474	
仮説平均との差異	0	
自由度	23	
t	2.556517532	
P(T<=t) 片側	0.008820643	
t 境界値 片側	1.713871528	
P(T<=t) 両側	0.017641286	
t 境界値 両側	2.06865761	

図7 パ・リーグ勝率の差の検定

セ・パともに帰無仮説を棄却する
 ⇒ホームの方がアウェイより勝率が高い

※ (再掲) 3. 研究内容

仮説

- I. **観衆因子**がホームアドバンテージに影響を与えている
⇒ホーム環境の方がファンの数や応援の熱が強いことが影響
- II. **ルール因子**がホームアドバンテージに影響を与えている
⇒ルール上、ホームチームが後攻になることが影響
- III. **習熟度因子**がホームアドバンテージに影響を与えている
⇒シーズンの中の約半分がホーム球場で行われることが影響

7-1. 仮説 I 観衆因子

分析手法・内容

ロジスティックス回帰分析

18-19年、20-21年の1試合ごとのデータを用いてリーグ別に分析

目的変数：勝敗関数（勝ち：1、負け：0）

説明変数：満席率（入場者数だとキャパの違いに左右されるため）

分析結果

どのデータにおいても統計的に有意な結果は得られなかった

		方程式中の変数					
		B	標準誤差	Wald	df	有意確率	Exp(B)
ステップ1 ^a	満席率	.894	.640	1.953	1	.162	2.445
	定数	-.700	.574	1.484	1	.223	.497

a. ステップ 1: 投入された変数 満席率

		方程式中の変数					
		B	標準誤差	Wald	df	有意確率	Exp(B)
ステップ1 ^a	満席率	.183	.388	.221	1	.638	1.200
	定数	-.068	.310	.048	1	.827	.934

a. ステップ 1: 投入された変数 満席率

		方程式中の変数					
		B	標準誤差	Wald	df	有意確率	Exp(B)
ステップ1 ^a	満席率	.894	.640	1.953	1	.162	2.445
	定数	-.700	.574	1.484	1	.223	.497

a. ステップ 1: 投入された変数 満席率

		方程式中の変数					
		B	標準誤差	Wald	df	有意確率	Exp(B)
ステップ1 ^a	満席率	.073	.600	.015	1	.903	1.076
	定数	-.140	.142	.968	1	.325	1.150

a. ステップ 1: 投入された変数 満席率

図8 18-19年セ・リーグ分析結果

図9 18-19年パ・リーグ分析結果

図10 20-21年セ・リーグ分析結果

図11 20-21年パ・リーグ分析結果

7-1. 仮説 I 観衆因子

分析手法・内容

相関分析

18-19年、20-21年の1試合ごとのデータ（計3310試合）を用いて

リーグ別に満席率と攻守各指標の関係を分析

⇒勝敗ではなく攻守指標への影響をはかるため

※入場者規制の関係で18-19年度と20-21年度を別々に分析

分析結果

全指標において無相関であることが分かった



ホーム環境の方がファンの数や応援の熱が強いことが影響しているという観衆因子における仮説 I は否定される

表 8 18-19セ・リーグ相関分析結果

Column_0	Column_1	correlation
ホームチーム得点	満席率	0.044168
アウェイチーム得点	満席率	0.015348
ホームチーム安打数	満席率	-0.000536
ホームチーム本塁打	満席率	0.069154
アウェイ安打数	満席率	0.003506
アウェイ本塁打	満席率	0.017483
ホーム打率	満席率	0.013006
ホーム出塁率	満席率	0.040482
ホーム長打率	満席率	0.041436
ホーム三振	満席率	0.055758
アウェイ打率	満席率	-0.010497
アウェイ出塁率	満席率	0.004009
アウェイ長打率	満席率	-0.005745
アウェイ三振	満席率	0.072504
ホーム防御率	満席率	0.001398
ホーム与四球率	満席率	0.061522
アウェイ防御率	満席率	0.048931
アウェイ与四球率	満席率	0.091973

7-2. 仮説Ⅱ ルール因子

分析手法・内容

ウィルコクソンの符号順位和検定

18-21年両リーグの地方球場勝率（ホーム主催）とアウェイ球場勝率

（ホームチームが後攻であるというルールの影響を測るパターンで分析）

帰無仮説：地方球場勝率（ホーム主催試合）とアウェイ球場勝率に差がない

対立仮説：地方球場勝率（ホーム主催試合）とアウェイ球場勝率に差がある

		地方球場勝率	アウェイ球場勝率	差	差の絶対値	順位
セ・リーグ	2018	0.5	0.476303318	0.023697	0.0236967	2
	2019	0.5	0.447368421	0.052632	0.0526316	4
	2020	0.25	0.422619048	-0.17262	0.172619	7
	2021	0.583333333	0.493472585	0.089861	0.0898607	5
パ・リーグ	2018	0.487804878	0.519047619	-0.03124	0.0312427	3
	2019	0.318181818	0.471291866	-0.15311	0.15311	6
	2020	0.666666667	0.424418605	0.242248	0.2422481	8
	2021	0.47826087	0.495978552	-0.01772	0.0177177	1

W	19
w	17
T（検定統計量）	17
n	8
限界値（有意水準 $\alpha=0.05, n=8$ で両側検定）	3

図12 ウィルコクソンの符号順位和検定 分析結果

検定統計量（T） > 限界値であるため、帰無仮説は棄却されない
 ⇒ 勝率に差は確認できないという仮説Ⅱを否定する結果になった
 （ルール因子はホームアドバンテージに影響していない）

7-3. 仮説Ⅲ 習熟度因子

分析手法・内容

t 検定

18-21年の【ホーム球場戦績】と【その他球場戦績】の各変数の年度別データを用いた
⇒年に40%程度使用する球場とその他の球場を比較することで

ホームアドバンテージにおける習熟度因子を測る

帰無仮説：ホーム球場戦績とその他球場戦績に差がない

対立仮説：ホーム球場戦績とその他球場戦績に差がある

扱った変数

勝率、打率、出塁率、長打率、三振数
防御率、与四球率、被打率、奪三振数
1試合当たりの得失点・本塁打数・被本塁打数

分析結果

有意水準を満たしたものが青で塗りつぶしたものである
またそれらは全て帰無仮説を棄却する結果になった

表9 習熟度因子t検定分析結果

扱った変数
勝率
打率
出塁率
長打率
三振数
防御率
与四球率
被打率
奪三振数
1試合当たりの得点
1試合当たりの失点
1試合当たり本塁打数
1試合当たりの被本塁打数

7-3. 仮説Ⅲ 習熟度因子

結果

勝率：ホーム球場の方が高い傾向にある

攻撃指標

出塁率：ホーム球場の方が高い傾向にある
三振数：ホーム球場の方が少ない傾向にある

守備指標

防御率：ホーム球場の方が低い傾向にある
与四球率：ホーム球場の方が低い傾向にある



攻撃においてはホーム球場の方が三振が少なく出塁率が高い傾向にあり、
より多くのチャンスを作っていることが考えられる
ただ打率や得点に繋がっているとは統計的に言えない
守備においてはホーム球場の方が与四球率、防御率が低く、
投手は無駄な走者を出さず、失点を防いでいることが分かる



以上のことがホーム球場の方が勝率が高い結果に繋がっていると読み取れる

8. 考察

仮説Ⅰ：観衆因子

統計的に有意な結果は得られなかった

⇒もちろん観衆の密度によって選手のモチベーションアップなど
心理的面の影響は考えられるが競技のルール上問題点はないため
チームやリーグ単位でデータを扱ったこの研究では
その影響は全体的には表れなかった
選手個人単位で見ると影響がある可能性は考えられる

仮説Ⅱ：ルール因子

統計的に有意な結果は得られなかった

⇒ルール因子の影響は試合展開によって異なる可能性がある為
全試合で見ると中々その影響はでないと考えられる
またサンプル数も少なく試合展開を絞った研究ができなかったため
この結果になったとも考えられる

8. 考察

仮説Ⅲ：習熟度因子

いくつかの攻守指標において統計的に有意な結果が得られた

➤ 中でも守備面の投手における指標において大きく影響している

理由としては球場ごとに異なるマウンドの高さや土質傾斜などが考えられる

その違いは、繊細な投手の感覚に強く影響することが考えられ、

分析を通しホーム球場において防御率や与四球率が低いことが明らかになり、それを証明したともいえる

9. まとめ

結論

ホームアドバンテージの要因は

“習熟度因子”で中でも投手において有利であることが要因

野球は点取りゲームであり、極端な話点が取られることがなければ負けることもなく勝率も上がる

その為、ホーム球場において防御率や与四球率が低いホームチームにアドバンテージがあると考えられる

9. まとめ

以上のことからNPBのチームに次のことを提案する

ホーム球場並びに
同リーグの他チーム球場のマウンドを
再現したブルペンを作ること

他球場の環境（マウンド）にもなれることで防御率を下げ、
勝率を伸ばし、リーグ優勝にも繋がる

観客動員数・チケット収入UPに繋がる

10. 今後の課題

- ルール因子などについては地方球場での試合数が他と比べると少なく、より多くのデータを扱うことができれば詳細な分析も可能であったと考えられる
- 今回は試合データを用いたホームアドバンテージの要因分析であったが、例えばホーム球場でホームチームは大浴場があることや、食堂のメニューの豊富さなど外的環境がホームアドバンテージに影響を与えていることも考えられる

参考文献

- 森和也, 「日本プロ野球における観客の有無による試合結果への影響についての研究」, 2022年
- 青山晃丈, 「プロ野球における勝率の分析」, 2015年
- 池侑弥, 「プロ野球における勝率に関する統計的検定」, 2015年
- 瀧上凱令, 「日本プロ野球におけるホームアドバンテージ」, 2013年
- 原田尚幸, 守能信次, 原田宗彦, 菊池秀夫
「ホームアドバンテージと観衆要因に関する研究」, 1996年
- 田代浩汰, 本田明生 「J1リーグにおけるホームアドバンテージの検証-通常試合と無観客試合の比較-」, 2020年
- 安部健太, 「慣れのホームアドバンテージへの影響の検討」, 2018年
- 笹川スポーツ財団

<https://www.ssf.or.jp/index.html> (最終閲覧日2022/7/13)

- 日本野球機構 統計データ

<https://npb.jp/statistics/> (最終閲覧日2022/11/10)

- 公益財団法人 全日本野球連盟

<https://jsbb.or.jp/> (最終閲覧日2022/7/12)

Appendix

◆ペナントレース

プロ野球の公式戦のことを意味する

約一年間（4月～10月）をかけてセ・リーグ、パ・リーグで行われる

◆コロナ禍の入場規制について

2020年のペナントレース開幕から2020年7月までは無観客試合を行いそれ以降は上限を設定しての有観客試合再開

2022年から制限を設けずに試合を行うようになった

◆平均値の差の検定(t検定)

t検定-帰無仮説が正しいと仮定した場合に統計量がt分布に従うことを利用する統計学的検定法

2つの条件・群の平均値を比較する統計的仮説検定の方法

◆ロジスティック回帰分析

ロジスティック回帰分析は、いくつかの要因（説明変数）から「2値の結果（目的変数）」が起こる確率を説明・予測することができる統計手法で、多変量解析の手法の1つ

◆出塁率

野球において打者を評価する指標の一つ。打者の打撃機会当たりの出塁割合を表す

◆与四球率

野球において投手を評価する指標の一つ。投手が一試合当たりの平均与四球数を表す

Appendix

◆ 相関分析

2つのデータにおける関係性の強さを分析する手法

相関係数 = x と y の共分散 / x の標準偏差 × y の標準偏差

◆ Wilcoxon（ウィルコクソン）の符号順位和検定

ノンパラメトリック検定のひとつで対応のある2群の差の検定に用いる。

対応のある2群のデータの差の絶対値を小さいものから順に並べ、順位付けを行う。次に、正の符号を持つ順位和の値から2群の差の有無を検定する。次に、が正の値の合計順位と負の値の合計順位を求め、2つの合計順位の小さい方を検定統計量とする。このを用いて両側検定のみ行い、検定統計量が統計数値表の有意水準5%あるいは1%の棄却限界値以下の場合、帰無仮説が棄却される。

◆ プロ野球の主な収入源

▶ チケット収入、放映権収入、グッズ販売収入、スポンサー収入

◆ NPB勝率 = 勝ち数 / (勝ち数 + 負け数)

◆ 満席率 = 観客動員数 / 球場のキャパシティ

Appendix

◆マウンド規定

直径5.4864cmの円形

高さは254mm

1フィートにつき1インチの傾きを付けなければならない

土質などに規定はなく球場ごとに異なる

◆習熟度 t 検定分析結果

勝率	ホーム球場勝率		その他球場勝率		出塁率	ホーム球場出塁率		その他球場出塁率		三振数	ホーム球場三振数		その他球場三振数	
	平均	分散	観測数	ピアソン相関		仮説平均との差異	自由度	t	P(T<=t) 片側		t境界値 片側	P(T<=t) 両側	t境界値 両側	
平均	0.534648685	0.467603984	0.21117814	0.205490975	平均	458.4166667	565.4375							
分散	0.007951582	0.004929817	0.000235266	0.000352892	分散	1721.652482	6937.230053							
観測数	48	48	観測数	48	観測数	48	48							
ピアソン相関	0.119708163		ピアソン相関	0.597646582	ピアソン相関	0.249427395								
仮説平均との差異	0		仮説平均との差異	0	仮説平均との差異	0								
自由度	47		自由度	47	自由度	47								
t	4.353793078		t	2.523744767	t	-8.903683742								
P(T<=t) 片側	3.59265E-05		P(T<=t) 片側	0.007525028	P(T<=t) 片側	5.91322E-12								
t境界値 片側	1.677926722		t境界値 片側	1.677926722	t境界値 片側	1.677926722								
P(T<=t) 両側	7.1853E-05		P(T<=t) 両側	0.015050055	P(T<=t) 両側	1.18264E-11								
t境界値 両側	2.011740514		t境界値 両側	2.011740514	t境界値 両側	2.011740514								

防御率	ホーム球場防御率		その他球場防御率		与四球率	ホーム球場与四球率		その他球場与四球率		
	平均	分散	観測数	ピアソン相関		仮説平均との差異	自由度	t	P(T<=t) 片側	t境界値 片側
平均	3.983735612	4.522078703	平均	3.895408389	4.418083533					
分散	0.506661773	1.682655137	分散	0.305013889	1.479890863					
観測数	48	48	観測数	48	48					
ピアソン相関	0.094958105		ピアソン相関	0.181145963						
仮説平均との差異	0		仮説平均との差異	0						
自由度	47		自由度	47						
t	-2.628174268		t	-2.916627979						
P(T<=t) 片側	0.005779937		P(T<=t) 片側	0.002704768						
t境界値 片側	1.677926722		t境界値 片側	1.677926722						
P(T<=t) 両側	0.011559875		P(T<=t) 両側	0.005409536						
t境界値 両側	2.011740514		t境界値 両側	2.011740514						