

読売ジャイアンツのセ・パ交流戦における勝敗に関する統計的分析

2016SS052 野口将史

指導教員：白石高章

1 はじめに

私は幼い頃から野球をしており、今では野球場に行き試合観戦をすることが趣味である。その中で毎年セ・パ交流戦でセ・リーグがパ・リーグに大きく負け越していることに注目した。そこで、セ・リーグで最も多くセ・パ交流戦を優勝したことがあり、通算勝率が5割を超えている「読売ジャイアンツ」(以後：巨人)を対象に優勝することができた2年分のデータと優勝できなかった5年分のデータを用いてパ・リーグ球団に対しての戦術、攻撃の特徴の比較を行い、また試合の勝敗を左右する要因は何かを統計的に分析することにした。

2 データについて

本研究では、巨人の2015年～2019年までを対象とした5年間のセ・パ交流戦の試合データ(以後：データA)と過去2度優勝した2012年と2014年の2年間のセ・パ交流戦の試合データ(以後：データB)を使用する。データは、文献[1],[2]から「勝敗(勝ち=1,負け=0)」「ホーム&ビジター(ホーム=1,ビジター=0)」「打数」「得点」「単打」「二塁打」「三塁打」「本塁打」「盗塁」「犠打」「犠飛」「四球」「死球」「三振」「失策」「被安打」「被本塁打」「与四死球」「奪三振」「失点」の20要素を用いる。

3 分析方法について

分析方法としてはロジスティック回帰分析、重回帰分析、クラスター分析を行った。分析は文献[3],[4]を参考に行った。重回帰分析では攻撃の変数と守備の変数で目的変数を変えた。さらに、ロジスティック回帰分析と重回帰分析で変数選択(ステップワイズ法)を行った。

4 ロジスティック回帰分析

データA, データBを用いて勝敗を目的関数とし、その他の変数を説明変数としてロジスティック回帰分析を行った。ステップワイズ法により表1の変数を説明変数に置いたモデルが最適となった。

表1 データA(左), データB(右)

変数	回帰係数	標準誤差	p値	回帰係数	標準誤差	p値
得点	0.600	0.065	0.001	0.411	0.083	0.001
犠打	-	-	-	0.218	0.085	0.014
犠飛	-	-	-	0.132	0.079	0.104
四球	-	-	-	0.155	0.086	0.080
失点	-0.545	0.065	0.001	-0.667	0.079	0.001

4.1 考察

データA, データBともに得点には正の相関、失点には負の相関が出た。このことから、勝敗には得点と失点が深く関わってくることを改めて理解できた。データAでは得点と失点しか相関関係が出なかったが、データBでは犠打、犠飛、四球に正の相関が出た。安打に相関関係は出ず、犠打、犠飛に正の相関が出たことから1つのアウトを犠牲にする戦術が勝利に大きく関わってくるのがわかる。また、四球に関しては安打以外で出塁することで試合の流れも大きく変わり、得点のチャンスも増えるのではないかと考える。

5 重回帰分析

5.1 攻撃時データを用いた重回帰分析

データA, データBを用いてそれぞれ得点を目的関数とし、得点に関係すると考えられる攻撃時の単打、二塁打、三塁打、本塁打、盗塁、犠打、犠飛、四球、死球、三振の10個のデータを説明変数とし、重回帰分析を行った。ステップワイズ法により表2の変数を説明変数に置いたモデルが最適となった。

表2 データA(左), データB(右)

変数	回帰係数	標準誤差	p値	回帰係数	標準誤差	p値
単打	0.435	0.068	0.001	0.609	0.099	0.001
二塁打	0.412	0.062	0.001	0.450	0.122	0.001
三塁打	-	-	-	-0.300	0.125	0.022
本塁打	0.608	0.062	0.001	-	-	-
盗塁	-	-	-	0.178	0.097	0.081
犠打	-	-	-	-	-	-
犠飛	0.164	0.068	0.018	-	-	-

5.2 考察

データAでは安打の中でも本塁打が最も高い相関関係が出ており、打撃力、長打力を活かし得点を重ねている。小技の相関もないことから、打って点を取りに行く打撃のチームということがわかった。一方、データBでは安打の中でも単打、二塁打に高い相関関係が出た結果となり、主に打線のつながりで得点を重ねている。また、盗塁にもわずかながら正の相関が出ており、優勝したシーズンは打撃力に加え、機動力も活かし得点を重ねるチームということがわかった。

5.3 守備時データを用いた重回帰分析

データA, データBを用いてそれぞれ失点を目的関数とし、失点に関係すると考えられる守備時の失策、被安打、被本塁打、与四死球、奪三振の5つのデータを説明変数とし、重回帰分析を行った。ステップワイズ法により表3の変数

を説明変数に置いたモデルが最適となった。

表 3 データ A(左), データ B(右)

変数	回帰係数	標準誤差	p 値	回帰係数	標準誤差	p 値
被安打	0.506	0.056	0.001	0.600	0.100	0.001
被本塁打	0.630	0.057	0.001	0.361	0.095	0.001
与四死球	0.137	0.057	0.019	0.138	0.099	0.171

5.4 考察

データ A では被本塁打に強い正の相関が出た。このことから、安打よりも圧倒的に本塁打で失点していることがわかる。一方、データ B では被安打に強い正の相関が出ており、主に単打、二塁打による連打によって失点していることがわかる。データ A, データ B ともに与四死球にも正の相関が出ており、相手に余計な得点のチャンスを与え、失点につながっていると考えられる。

6 クラスタ分析

各年の試合をクラスタ分析で群分けを行い、どのような試合の特徴が多いと勝ちにつながるかを分析、考察を行った。

6.1 2014 年考察

各年の群分けの中できれいに群分けすることができた 2014 年を対象に考察を行う。

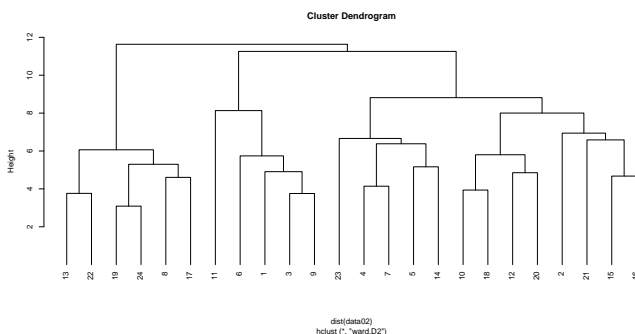


図 1 2014 年デンドログラム (ウォード法)

2014 年は得点、安打が少なく敗戦した試合が 6 試合しかなく、その他の試合はパ・リーグ球団に圧倒的に勝利した試合が多い。安打の中でも単打、二塁打、本塁打の数も多く、また盗塁と四球の数も多い特徴がある。それらに加え僅差での試合を勝ち切ることも多いことから投打がかみ合い、打撃力、機動力、選球眼を活かし勝利を重ねたシーズンといえる。

6.2 2012 年, 2015 年~2019 年考察

好成績を収めることができた 2012 年と 2019 年は 2014 年と同様に単打、盗塁、四球が多い特徴がある。優勝することができた 2012 年と 2014 年は僅差での試合が多く勝率も高いが、順位が 3 位だった 2019 年は僅差での試合は少

なく勝率もそれほど高くないことから接戦をものにするには優勝することに必要不可欠ということがいえる。一方、2015 年~2018 年の試合は三振が多く、失点が多い試合が目立つ。大差をつけて勝つという試合は少なく、得点は平均的で僅差での試合が多い。僅差での試合の勝率が低く勝ちきれないことからパ・リーグ球団に大きく負け越していると考えられる。

7 考察

優勝することができた巨人と優勝できなかった巨人の大きな違いは、攻撃スタイル、投手力、僅差の試合の勝率である。重回帰分析、ロジスティック回帰分析の結果から優勝できなかった巨人は小技は使わず、打って勝つ攻撃スタイルということがわかる。野球のセオリーと呼ばれる犠打の数は少なく、得点への相関もほとんどないという結果からまさに打線が売りのチームといえる。失点に関しては主に本塁打で失点しており、まさにパ・リーグの強力打線に捕まり、相手に勢いを与えてしまっている。一方、優勝することができた巨人は打撃力に加え、小技、機動力も活かし得点を重ねるチームといえる。単打、二塁打に強い正の相関が出ていることから一発は狙わず、打線のつながりを得意とするチームということがわかる。また、打つことだけではなく、犠打、犠飛、四球、盗塁など多くの戦術を駆使して得点を重ねており、まさに理想の攻撃スタイルを持っている。失点に関しては主に安打で失点していることから、大量失点につながらず最小失点で抑え、少ない得点で勝利する試合も目立つ。クラスタ分析の結果から、好成績を収めたシーズンはチャンスに強く、僅差での試合での勝率が高い。一方、下位成績に沈んだシーズンはチャンスに弱く、僅差での試合での勝率が低いことから勝利にはチャンスに強い勝負強さと接戦をものにする粘り強さが必要といえる。

8 終わりに

チームには「強力打線」や「小技スタイル」と種類があるが、優勝した年の巨人はその両方を兼ねそろえている万能型のチームということがわかった。近年の野球では 2 番打者にパワーヒッターを置くことが多くなってきているが、やはり犠打、盗塁などでランナーを確実に進塁させる攻撃のほうで得点圏にランナーがいるというプレッシャーを相手にかけることができ、より得点のチャンスも増え勝利に近づくと考える。失点に関しては常に本塁打を警戒し、相性の悪い打者とは勝負を避けるなどの対策が必要と感じた。

参考文献

- [1] 『プロ野球データ Freak』
<https://basaball-data.com/>
- [2] 『NPB.jp 日本野球機構』 <http://npb.jp/>
- [3] 金 明哲: 『R によるデータサイエンス データ解析の基礎から最新手法まで』 森北出版, 東京, 2007
- [4] 青木 繁伸: 『R による統計解析』 オーム社, 東京, 2009