ソフトバンクホークスに関する統計解析

学生番号 2018SS027 梶浦翔斗

指導教員:白石高章

1　はじめに

私は高校まで野球をやっており, プロ野球の結果や情報を日々見ることが趣味である. その中で, 私は近年ソフトバンクホークス (以後:ソフトバンク) が他チームを圧倒し, 多くのリーグ優勝に輝いていることに注目した.

しかし, ソフトバンクでもなかなか結果を出せずに最下位に沈んだ時期もあり, ソフトバンクがいかにしてここまでの常勝軍団になっていった要因は何か. また, ソフトバンクは今シーズン 8 年ぶりに B クラスに沈んでしまった要因知るべく各年で投球内容, 攻撃の比較を行っていくことにした.

2　データ

本研究では, 11 年間で 6 度のリーグ優勝に輝いた 2010 年〜2020 年を常勝期とし,4 年間で 1 度も優勝することが出来ず最下位に沈んだこともある 2006〜2009 年を衰退期とする. ソフトバンクが 2006 年〜2021 年までの 16 年間 (2387 試合) の公式戦データ を使用する.

データは日本プロ野球機構 [1] から「x1:得点」「x2:単打」「x3:二塁打」「x4:三塁打」「x5:本塁打」「x6:盗塁」「x7:四球」「x8:故意四球（打)」「x9:死球」「x10:併殺打」「x11:犠打・犠飛」「x12:三振」「x13:被安打 」「14:被本塁打」「x15:奪三振」「x16:与四球」「x17:与死球」「x18:故意四球 (守)」「x19:自責点」「x20:失点」「x21:セーブ」「x22:ホールド」「x23:完投」「x24:完封」の 24 要素を用いる. 分析方法は, 重回帰分析, クラスター分析, 得点, 得失点差による勝率の比較を行った. 試合のデータは, 交流戦を含めたレギュラーシーズンの試合のみとする.

3　重回帰分析

常勝期, 衰退期で得点, 失点を目的関数として, 得点に関する 11 要素と失点に関する 11 要素を説明関数として重回帰分析を行った.

3.1　攻撃時の重回帰分析

ステップワイズ法を行った結果表 1 の変数を説明変数に置くことが最適のモデルとなった. p 値が 0.05 以下で信頼度が高い説明変数の中から, 回帰係数の数値が高い順に並べると, 常勝期は本塁打, 二塁打, 併殺打, 故意四球 (打) となる. このことから, 常勝期は本塁打などの長打を中心とした攻撃をしていたことが考えられる. また, 衰退期も同様に並べると, 三塁打, 二塁打, 本塁打, 四球, 犠打, 犠飛, 三振となる. このことから, 衰退期は打撃だけではなく, 小技などを生かして得点に繋げていたことが考えられる.

表 1　常勝期 (左), 衰退期 (右)

説明変数	回帰係数	標準誤差	p 値	回帰係数	標準誤差	p 値
単打	-	-	–	0.08	0.04	0.08
二塁打	0.91	0.27	0.01	0.78	0.28	0.03
三塁打	-	-	-	3.19	1.09	0.03
本塁打	1.34	0.19	0.01	0.74	0.28	0.04
盗塁	-0.24	0.20	0.26	-	-	-
四球	-	-	-	0.72	0.13	0.01
故意四球 (打)	-3.01	1.06	0.02	-1.85	1.57	0.28
死球	-	-	-	2.60	1.19	0.07
併殺打	-1.24	0.46	0.03	-	-	-
犠打, 犠飛	0.57	0.30	0.10	0.57	0.20	0.03
三振	-0.20	0.10	0.08	-0.36	0.09	0.01

3.2　守備時の重回帰分析

ステップワイズ法を行った結果表 2 の変数を説明変数に置くことが最適のモデルとなった. 攻撃時と同様に並びかえると常勝期は, 自責点, 被本塁打, 与四球, 与死球, 完封となり, 衰退期は, 自責点, 被本塁打, ホールド, 故意四球 (守), 完投となった. この結果から, 常勝期は与四球が正の値が出ており, 四球などから失点に多く繋がっていたと考えられる. 衰退期はホールドの項目で負の値を出せていることから, 安定したリリーフ陣が整っていたと考えられる.

表 2　常勝期 (左), 衰退期 (右)

説明変数	回帰係数	標準誤差	p 値	回帰係数	標準誤差	p 値
被安打	-	-	–	0.06	0.03	0.14
被本塁打	0.55	0.06	0.01	0.51	0.12	0.01
奪三振	0.06	0.03	0.11	-	-	-
与四球	0.33	0.08	0.01	-0.07	0.04	0.11
与死球	-0.77	0.31	0.05	-0.25	0.18	0.22
故意四球 (守)	0.47	0.30	0.19	-0.72	0.32	0.05
自責点	0.98	0.05	0.01	1.25	0.08	0.01
セーブ	0.52	0.29	0.14	-	-	-
ホールド	-0.34	0.15	0.07	-0.29	0.11	0.04
完投	1.11	0.61	0.12	-2.73	0.54	0.01
完封	-1.37	0.54	0.05	-	-	-

4　クラスター分析

各年 (2006 年〜2021 年) をクラスター分析で群分けを行い, 攻撃時, 守備時でお互いにどのような特徴があるのか分析, 考察を行った. (2022-i) 年を第 i 番目 (i=1,⋯,16) とする.

4.1　攻撃時

ウォード法を用いたクラスター分析で 3 つの群に分け考察する.

第 1 群 (1(2021 年)〜5(2017 年)):本塁打が多く, 単打が少ない群である. 本塁打に関しては 1 番少ない第三群と比べると 1.5 倍もの差がついており, 長打を活かして得点に繋げているシーズンだったことが考えられる.

第 2 群 (6(2016 年)〜7(2015 年),9(2013 年)):単打, 四球, 得点が多い群である. この群は単打, 得点の数値が非常に高く, 四球も高いことから安打でのチャンスメイクはもちろん, 選球眼にも優れており, 多くの得点につながったシーズンだったと考えられる.

第 3 群 (8(2014 年),10(2012 年)〜16(2006 年)): 得点と本塁打が少なく, 盗塁が 多い群である. この群は本塁打が非常に低く, 単打は多いことから, 単打などで出塁してからの機動力を絡めた攻撃をしていたがチャンスでの安打が出ず, 得点になかなかつながらなかったシーズンだったと考えられる.

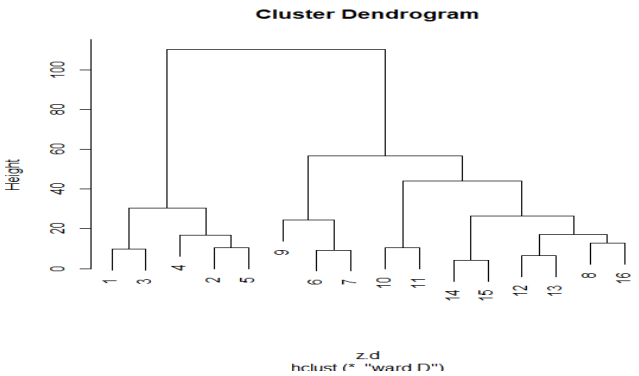

図 1　攻撃時デンドログラム (ウォード法)

4.2　守備時

ウォード法を用いたクラスター分析で 3 つの群に分け考察する.

第 1 群 (1(2021 年)〜9(2013 年),12(2010 年)): 被本塁打, 奪三振, セーブ, ホールドが多く, 完投が少ない群である. 近年は投手の平均球速のアップによるパワーピッチャーの増加で球の勢いが強いからこそ被本塁打や奪三振が多くなっていると考える. また, セーブやホールドが非常に多いことは近年の投手陣は分業制が主流となってきていることから, 完投する投手の減少につながっていると考えられる.

第 2 群 (10(2012 年)〜11(2011 年)):被本塁打, 与四球, 失点も少ない群である. この群はほかの群と比べて大半の項目で良い成績を出しており, 特に安定した投手陣だったシーズンだと考えられる.

第 3 群 (13(2009 年)〜16(2006 年)) 被安打, 与四球, 失点が多く, セーブ, ホールドが少ない群である. この群はほかの群と比べて, 被安打, 与四球共に多く, 多くの出塁機会を与えてしまい多くの失点につながってしまったシーズンだと考えられる.

5　得失点差による勝率の分析

得失点差による勝率を求め, 棒グラフ化して, 常勝期の年間順位 1 位だった 2020 年と衰退期の年間順位 6 位だった 2008 年と年間順位 4 位だった 2021 年を比較する.

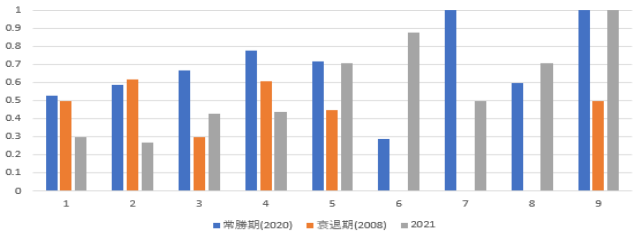

図 2　得失点差による勝率

得失点差で勝率を求めたところ, 常勝期はどのような点差がついていても安定して高い勝率を残しており, それに比べ, 衰退期は接戦での勝率は高いが点差がついた試合では勝率が低く, 2021 年は, 大差のついた試合での勝率は高いが, 接戦での勝率は低い. この結果から衰退期は圧倒的な大差で負ける展開が多く, 2021 年は接戦で勝ち切ることができない展開が多いことが考えられる.

6　おわりに

本研究を通して, ソフトバンクの戦い方は大きな変化を遂げていると感じた. 攻撃に関しては日本人らしい機動力, 小技を活かした野球から, メジャーリーグのような長打で大量得点を取る野球になっており, 投手陣に関してはより多くのリリーフ陣を起用し, 短いイニングを継投して少ない失点で逃げ切るスタイルへと変化していた. ソフトバンクは衰退期を打開するため, どこの球団よりも早く三軍制の導入やウェイトトレーニングの強化などの多くの取り組みを行い戦い方を大きく変化してきた結果, 他チームを圧倒する常勝軍団を築き上げることができたと考えられる. また, 今シーズン結果が残せなかった要因としては, 今シーズンはコロナ渦の影響で 9 イニング制がとられており, 他チームとの戦力差が出にくかったことがこの結果に繋がった一番の要因だと考える. 今後もソフトバンクが勝利のためにどのように変化を遂げていくのか注目していきたい.

参考文献

[1]『日本野球機構』https://npb.jp/(2021 年 11 月閲覧)

[2] 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 年

[3] 櫻井緑風:『中日ドラゴンズの黄金期に関する統計的分析』 2020 年度 南山大学理工学部システム数理学科 卒業論文.