

# プロ野球の試合時間に関する統計的分析

2010SE132 村山和也

指導教員：松田真一

## 1 はじめに

私は幼い頃から野球をしており、プロ野球や高校野球に非常に興味を持っている。そこで、近年プロ野球の試合時間が長く見ていられないと言う事が理由で野球を見なくなる、野球から離れていく人が増えているという事に興味を持った。また長時間、球場やドームの電気を使う事は地球温暖化につながるといった事から日本プロ野球連盟も試合を短縮させる動きがある。この野球離れを理解するために、実際の試合の内容といった視点から理解してみようと考えこのテーマを選んだ。私の研究は、試合時間に大きな影響を与えている要因は何なのかを分析するといったものである。統計的分析方法としては重回帰分析を用いる。

## 2 データについて

データは日本野球機構オフィシャルサイト(日本野球機構オフィシャルサイト [3] 参照)の2012年データを用いる。試合時間に影響を及ぼすであろう要因としてエラー数、得点数、安打数、三振数、四球数、死球数、登板投手数、代走数、代打数、イニング数といった10個の要因を用いる。これらのどの要因が試合時間に大きな影響を与えているのか理解する。セ・リーグ、パ・リーグ、交流戦でどういった違いが出てくるのか興味深いので「交流戦かそうでないか」、「セ・リーグかパ・リーグか」といった要因を加えた。また球場、ホームでの試合、アウェイでの試合では試合時間に関係があるのかを調べるために球場名を入れた。また、大リーグと日本プロ野球との試合時間の違いを調べるために大リーグのチームであるヤンキースの2013年の試合結果を用いた。(MLBオフィシャルサイト [4] 参照)データの中にイニング数があるが、9回の表で試合が終われば8.5とし、9回の裏で試合が終われば9とした。

## 3 解析の流れ

1. 試合時間に影響を及ぼしそうな要因を挙げる。
2. それらについてstepを用いて要因候補を絞る。
3. 絞った要因候補をVIF(分散拡大要因)にかけ、値が10以上の要因があればそれを除き2に戻り、値が10以上の要因がなければそれを要因と決定する。
4. それらの要因を用いて重回帰分析を行い、各要因の意味づけを行う。ここまでの工程を8パターンすべて行う。
  - セ・リーグ前半戦
  - パ・リーグ前半戦
  - 交流戦
  - セ・リーグ後半戦
  - パ・リーグ後半戦
  - セ・リーグ全体

- パ・リーグ全体
- セ・パ・リーグ全体

5. 全8パターンの結果より総合的に判断し、試合時間の算出の際に使用する要因を決定する。

## 4 重回帰分析

セ・リーグ前半戦、パ・リーグ前半戦、交流戦、セ・リーグ後半戦、パ・リーグ後半戦、セ・リーグ全体、パ・リーグ全体、セ・パ全体の全8パターンの分析を行い、「球場」、「ホーム」、「アウェイ」は試合時間に影響を及ぼす可能性が低い事が分かった。試合時間に影響を及ぼす可能性がある要因としてエラー数、得点数、安打数、三振数、四球数、死球数、登板投手数、代走数、代打数、イニング数の10個に決定し、重回帰分析を行った。また残差分析を行ったところ正規性があり、外れ値がなくデータに誤りがないと分かった。VIFを調べたところ、10以下で問題がなかった。

表1 10個の要因に関する重回帰分析(セ・パ全体)

	推定値	標準誤差	t値	P値
定数項	3.5377	7.0049	0.505	0.6136
エラー数	1.4335	0.4043	3.546	0.0004
得点数	-1.1439	0.1800	-6.354	0.0000
安打数	2.4173	0.1464	16.515	0.0000
三振数	0.9380	0.1116	8.406	0.0000
四球数	3.0445	0.1860	16.370	0.0000
死球数	3.3538	0.4824	6.952	0.0000
登板投手数	3.3195	0.2477	13.403	0.0000
代走数	2.6346	0.3944	6.680	0.0000
代打数	0.1594	0.2467	0.646	0.5183
イニング数	10.6244	0.8131	13.066	0.0000

- 決定係数 0.7478
- 自由度調整済み決定係数 0.7448

## 5 試合時間の予想計算式 - 日本プロ野球

「エラー数」、「得点数」、「安打数」、「三振数」、「四球数」、「死球数」、「登板投手数」、「代走数」、「代打数」、「イニング数」の10個の重回帰分析の結果は表1のようになった。表1より算出される試合時間予想値はこの計算式から求められる。

$$\begin{aligned} \text{[試合時間予想値]} = & 3.5377 + 1.4335 \times \text{[エラー数]} - \\ & 1.1439 \times \text{[得点数]} + 2.4173 \times \text{[安打数]} + 0.9380 \times \text{[三振数]} \\ & + 3.0445 \times \text{[四球数]} + 3.3538 \times \text{[死球数]} + 3.3195 \times \\ & \text{[登板投手数]} + 2.6346 \times \text{[代走数]} + 0.1594 \times \text{[代打数]} + \\ & 10.6244 \times \text{[イニング数]} \end{aligned}$$

## 5.1 コントロールの良い投手 (日本野球)

コントロールに関わってくる変数として「四球数」と「死球数」がある。1 試合平均の四球数, 死球数は 5.372 個, 0.693 個である。死球数は 1 試合に 1 個出るか出ないかである。試合時間予想の計算式から四球数を 1 試合 1 個にする事で平均試合時間よりも 12 分短縮することが出来る。コントロールの良い投手を登板させることで試合時間の短縮になると考えられる。

## 5.2 登板投手の少ない試合 (日本野球)

球数制限のない日本の野球で, 登板投手の多い試合は試合時間が長くなる傾向がある。その理由として投手交代が行われるのは, 投手が打たれてしまい失点をした時が多く, 安打の本数が多い時と考えられる。投手戦のように得点の少ない試合は投手が最小に安打, 得点を抑えているため投手交代も少なく, 結果的に試合時間は短くなる。

## 6 大リーグについて

日本とアメリカの野球の試合時間についての違いを見つけるために大リーグの試合を集めた。大リーグの試合すべてを集めるのではなく, ニューヨークヤンキースの試合 40 試合を集めて分析を行った。データについては MLB オフィシャルサイトから集め, 球場名, ホーム, アウェイの以外の「エラー数」, 「得点数」, 「安打数」, 「三振数」, 「四球数」, 「死球数」, 「登板投手数」, 「代走数」, 「代打数」, 「イニング数」の 10 個で分析を行った。(MLB オフィシャルサイト [4] 参照) 10 個の変数を step 1 かけ重回帰分析を行ったところ, step 2 を行う前よりも step 3 後の変数の方が, 自由度調整済み決定係数が高かったため, その時に残った「得点数」, 「三振数」, 「四球数」, 「死球数」, 「登板投手数」, 「代打数」の 6 個を試合時間に影響を及ぼす可能性がある要因として決定した。その結果を以下に示す。また残差分析を行ったところ正規性があり, 外れ値がなくデータに誤りがないと分かった。VIF を調べたところ, 10 以下で問題がなかった。

表 2 step 後の重回帰分析結果 - 大リーグ

	推定値	標準誤差	t 値	P 値
定数項	132.7501	5.6265	23.586	0.0000
得点数	2.7883	0.4275	6.522	0.0000
三振数	1.2014	0.3004	3.999	0.0003
四球数	1.7972	0.4060	4.426	0.0001
死球数	-4.5326	1.2856	-3.526	0.0012
登板投手数	1.5093	0.6851	2.203	0.0349
代打数	-3.2352	0.8076	-4.006	0.0003

- 決定係数 0.8686
- 自由度調整済み決定係数 0.8439

## 6.1 試合時間の予想 - 大リーグ

表 2 より算出される試合時間予想値はこの計算式から求められる。

- 試合時間の予想計算式 - 大リーグ

$$\begin{aligned} \cdot [\text{試合時間予想値}] = & 132.7501 + 2.7883 \times [\text{得点数}] + \\ & 1.2014 \times [\text{三振数}] + 1.7972 \times [\text{四球数}] - 4.5326 \times [\text{死球数}] \\ & + 1.5093 \times [\text{登板投手人数}] - 3.2352 \times [\text{代打数}] \end{aligned}$$

## 6.2 得点数を入れるほど試合時間は長くなる (大リーグ)

表 1, 表 2 より日本のプロ野球ではマイナスに作用していた「得点数」がプラスに作用している。日本は得点差がつくと淡白な攻撃をし試合時間が早まるといったことが考えられたが大リーグでは, 得点差がついても打ち続ける, またホームランが途切れないといったことが考えられる。それを調べるために, 5 点差以上の点差がついた試合を調べてみたところ, 日本のプロ野球が 60 試合中 8 試合だったのに対し, 大リーグでは 60 試合中 18 試合もあった。これは大リーグが日本よりも大量の点差をつけた試合が多いことを表しており, 打ち続ける傾向があると考えられる。

## 7 まとめ

試合時間について日本のプロ野球, 大リーグについて多くのデータを集め, 解析を行った。その中で分かった事として, 当たり前ではあるがプレー中におこる出来事は何かしら試合時間に関係があるということである。四球や死球, エラーなど選手の意識によって変えられるものは減らしていき, 攻守交替の時に高校球児の様に, 全力で交代したり, 審判の指示には従い遅延となる行為を減らすなど, 選手が試合時間を短縮しようとする事だと感じた。

## 8 おわりに

今回の研究を通して, 改めてプロ野球というものの面白さを感じられ, 野球離れをしていく人達を食い止めるためにも, だらだらした試合をせず短い試合でファンを楽しませるものになればよいと感じた。

## 参考文献

- [1] 中澤港:「Rによる統計解析の基礎」. ピアソン・エデュケーション, 2003.
- [2] 山田剛史・杉澤武俊・村井潤一郎:「Rによるやさしい統計学」. オーム社, 2008.
- [3] 日本野球機構オフィシャルサイト:  
<http://www.npb.or.jp/>
- [4] MLB オフィシャルサイト:  
<http://mlb.mlb.com/home>
- [5] 荒木考治:「RとR コマンドーではじめる多変量解析」. 日科技連出版社, 2007.
- [6] 野球情報サイト:  
<http://baseballmonster.nobody.jp/major/difference.html>