

# 野球における戦略評価方法の研究

データ科学研究系 計算機統計グループ  
客員教授 山口 和範 (立教大学)

## 1 はじめに

### 1.1 野球に関する統計分析

スポーツにおける統計の利用は古い歴史を持ち、世界の統計関連の学会ではスポーツ統計のセクションを持つところがある。日本においても昨年度日本統計学会にスポーツ統計分科会が設置され、その活動に注目が集まっている。野球は米国で誕生したスポーツであり、米国野球学会 (Society for American Baseball Research : SABR) が1971年に設立され、野球の経営や歴史さらには統計分析などのセクションに分かれ研究者のみならず野球愛好家も含めて活発な活動がなされている。この学会での統計分析に関する活動の成果は、この学会のSABRという頭文字からSABRmetricsと名付けられ、現在のメジャーリーグでの戦略や選手評価ならに選手リクルートをはじめとするチーム経営の常識となりつつある(たとえば、Lewis 2003 参照)。また、近年はFitts(2007)などをはじめとし、日本のプロ野球の戦略評価に関する論文も見受けられる。この学会活動における統計分析活動が隆盛となった1つの理由は、メジャーリーグでの実践の種々のデータが公開され、その解析が比較的自由に行えたことである。このような状況は、マーケティングリサーチにおけるPOSデータの公開が米国で行われ、その後多様な統計モデルの発展がみられたことと共通していると思われる。ただ、野球における試合結果のデータもPOSデータも一種の観察研究によるデータの収集であり、そこでの戦略評価においては交絡要因の考慮が必要となる。

### 1.2 犠牲バントに関する統計分析

野球において、アウトカウントを増やしてもランナーを1つだけ進める作戦である「犠牲バント」については、野球における戦略研究の対象として最も注目を浴びている内容である。この「犠牲バント」については、アメリカのメジャーリーグでの実データを用いた分析が、数多くなされており、その結果としてはAlbert and Bennett(2003)、Tango et al. (2007)、Thorn and Palmer (1985)などが有名である。議論の一例を紹介しよう。Albert and Bennett(2003)では、Lindsey(1963)の集計やメジャーリーグの2002年シーズンの集計から無死一塁での得点確率や平均得点数と一死二塁での得点確率と平均得点数を比較している。ここでは、得点確率や平均得点数が共に無死一塁より一死二塁の方が下がることが示されている。Tango et al. (2007)では、同様の分析に加え、勝敗を目的変数とした分析が行われている。たとえば、7回の裏で1点ビハインドの場合の無死一塁での勝利確率の推定値が0.413で、一死二塁の場合が0.403であることが例示されている。このことから、必ずしも「犠牲バント」は得策であるとはいきれないと結論付けている。もちろん、このような分析が一定の条件が与えられたときの作戦の選択とは直接結びつかないことは明らかで、各試合中での与えられた条件下での「犠牲バント」の評価が必要となる。野球の実戦で活用可能なためには、多様な状況を考慮しなければならず、そのような中での勝利または得点確率に対する「犠牲バント」の寄与を評価する一般モデルを構築することは非常に困難であると考えられる。そこで、今回の報告では、「犠牲バント」が顕著に有効または無効な条件を探索し、その信頼性を評価することを試みる。

## 2 決定木による条件選択

### 2.1 データと分析手順

日本のプロ野球公式戦での試合記録データに関する分析を行う。まず条件抽出のため2005年シーズンの公式戦全試合に関するデータを使用する。具体的には、全試合の全インニングごとのデータに関し、先頭打者が出塁して無死一塁の状況となったインニングを抽出し、その無死一塁での状況に関する情報および犠牲バントの有無、最終的なそのインニングでの得点の有無およびそのインニングでの得点を集計する。なおここでの分析対象とする状況は、先頭打者が出塁して無死一塁になった場面のみとし、本塁打などで得点が入った後に無死一塁の状況が発生しても分析対象の状況とはしていない。2005年でのデータで、対象となるケースは5803件であった。ここでは、犠牲バントの有効性を検証するため、決定木を使用し有効となる場面の抽出を行い、2008年度のデータでその抽出結果に関する有意性の検証を行う。

使用する変数としては、インニング、そのインニングまでの攻撃側および守備側の得点、打者打率や次打者打率など攻撃側の打力に関する指標、走者の走塁力に関する指標（たとえば盗塁数など）、さらに、守備側の守備力として投手力の指標や捕手の盗塁阻止率などを説明変数とし、目的変数としては得点の有無と得点数を用いた。以上のような変数の設定で決定木解析を行い、出力のなかから犠牲バントの有無で枝分かれする部分を抜き出し、その枝分かれ前の条件と枝分かれ後の得点確率（得点数の場合は得点数）の差を検討する。ここでは考慮すべき共変量をすべて取り込み十分な深さまで掘り下げて分析することで、交絡の影響を全てとはいえないまでも、可能な限り考慮（排除）し、犠牲バントが有効な、または逆効果の具体的場面を探索することが可能であると考えられる。

なお、2005年度データでの有意性の評価は不可能であるため、その評価には2008年度データを使用する。

### 2.2 分析結果

詳細な分析結果については当日報告する。

## 参考文献

Albert, J. and Bennett, J.(2003). *Curve Ball: Baseball, Statistics, and the Role of Chance in the Game*, Springer, New York.

Fitts, R. K.(2007).The Evolution of Japanese Baseball Strategy, *The Baseball Research Journal*, **36**, 61-67

Lewis, M.(2003). *Moneyball: The Art of Winning an Unfair Game*, W W Norton & Co Inc., New York.

Lindsey, G. R. (1963). An Investigation of Strategies in Baseball, *Operations Research*, **11** 477-501.

Tango, T., Lichtman, M. and Dolphin, A.(2007). *The Book: Playing the Percentages in Baseball*, Potomac Books, Virginia.

Thorn, J. and Palmer, P.(1985), *The Hidden Game of Baseball*, Doubleday, New York.