

Title	デッドボールが打者に与える影響度測定：デッドボール1球の価値とは
Sub Title	
Author	常松, 一喜(Tsunematsu, Kazuki) 林, 高樹(Hayashi, Takaki)
Publisher	慶應義塾大学大学院経営管理研究科
Publication year	2016
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2016年度経営学 第3182号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40003001-00002016-3182

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

慶應義塾大学大学院経営管理研究科修士課程

学位論文（ 2016 年度）

論文題名

デッドボールが打者に与える影響度測定
—デッドボール1球の価値とは—

主 査	林 高樹
副 査	河野 宏和
副 査	林 洋一郎
副 査	

氏 名	常松 一喜
-----	-------

論文要旨

所属ゼミ	林高樹 研究会	氏名	常松一喜
(論文題名)			
デッドボールが打者に与える影響度測定 —デッドボール 1 球の価値とは—			
(内容の要旨)			
<p>日本プロ野球の年間デッドボール数は約 600 近くあり、デッドボールによって故障する、あるいは調子を崩す選手は少なからず存在している。また、投手がデッドボールを投げる原因として①アクシデント②戦略的デッドボールの2つが考えられる。このようにデッドボールは打者に対して悪影響があるのにもかかわらず、近年日本プロ野球において 1 試合あたりのデッドボールの数が増えているという問題がある。そこでこの問題を解消すべく、2つの仮説検証を基にデッドボールの減少につながるペナルティの設計を行う。</p>			
【仮説①検証】			
<p>実際にデッドボールがどのように打者に影響を与えているのかを、ロジスティック回帰モデルを用いて定量的に測定した。その結果、デッドボールによって打者の出塁数は 0.049 回減少することが明らかになった。</p>			
【仮説②検証】			
<p>仮説①で出た結果を金銭的に評価すべく、出塁数によって打者の翌年度推定年俵にどのような影響があるのかを、重回帰モデルを用いて推定した。その結果デッドボール 1 球の価値は翌年度推定年俵の-0.017%と算出できた。</p>			
<p>以上の仮説検証の結果から、デッドボールによる打者の被害額は日本プロ野球全体として年間約 400 万円程度であると算出し以下の 3 つの点を踏まえたペナルティを設計した。</p>			
①ペナルティを支払うのは投手の所属球団			
②ペナルティの支払先は NPB (日本プロ野球機構)			
③NPB 主催のデッドボールの危険性を伝える講演会・野球教室の運営			

目次

1.序論	4
1.1 はじめに	4
1.2 研究背景	5
1.3 研究意義・目的	9
2.仮説設定	10
2.1 仮説	10
2.2 先行事例	11
2.2 データ収集	11
3. 仮説①検証	12
3.1 データ概要	12
3.2 ロジスティック回帰分析	16
3.3 仮説①結果考察	19
4. 仮説②検証	20
4.1 データ概要	20
4.2 重回帰モデル	25
4.3 仮説②結果考察	26
5. 総括	27
5.1 考察のまとめ	27
5.2 課題解決案	28
6. 今後の課題	29
7. 参考文献	30
8. 謝辞	31
付録 Rのコード	32

1.序論

1.1 はじめに

2011年、ブラットピッド主演の「マネーボール」という映画がヒットし、日本でも話題になった。この映画は資金力が乏しい球団が、セイバーメトリクスと呼ばれる、野球に関する統計データ分析を用いたチーム編成を行うことで、リーグトップクラスの成績をあげたオークランド・アスレチックスのビリービーンGMをモデルとしており、内容もさることながら「セイバーメトリクス」という野球の新たな切り口を世に浸透させた。このように野球を選手データから分析し、チーム編成・管理やゲーム戦略・戦術に用いる手法は近年、選手・監督・球団経営者のみならず野球ファンや、それ以外の人々にも認識されつつある。

このようにセイバーメトリクスはチーム編成・管理やゲーム戦略・戦術に活用されている一方で、セイバーメトリクスを用いて、野球という1スポーツの抱えている課題を発見して、課題を解決していくといったところには着手されてきてはいない。そこで本研究では、日本プロ野球界が抱えている課題を探索し、課題を解決する手段として統計的手法を用いた仮説検定と課題解決案について論じていく。

1.2 研究背景

筆者が、日本プロ野球が抱えている課題を探すにあたって、実際の試合や専門の雑誌だけではなく、野球をテーマとした漫画やプロ野球ファンの考えなど、幅広く見聞を広げてきた。その中で課題発見の一助となったものの一つに、MAJORという漫画がある。この漫画では、主人公の父親がデッドボールを受けて死んでしまい、また、主人公自身も投手としてデッドボールを投げてしまいイップスを発症し投球に支障をきたしている。これは漫画の話だからと一笑できるものではなく、現実にも逝去するまではいかなくとも、選手生命を奪われる選手や、デッドボール後本来のバッティングの調子を崩す打者はこれまでも少なからず存在している。例えば、広島東洋カープに在籍していた前田智徳は、2013年デッドボールによって左腕を骨折し、その年引退した。[1] また、2016年に2年連続トリプルスリーを達成した東京ヤクルトスワローズの山田哲人も、2016年ペナントレース終盤戦でデッドボールを受けてその後のバッティングの調子を落としていた。[2] このような実例からもデッドボールは打者に対して負の影響を与えることは明確であるが、デッドボールは実際に投手・野手それぞれにどのような影響を与えているのだろうか。投手側から見たデッドボールの影響を踏まえてまとめてみる。

デッドボールに関して、日本プロ野球もその危険性を認識しており、2000年以降デッドボール数を減らす施策として、頭部デッドボールに関しては、審判に裁量が委ねられるものの日本プロ野球規則により退場処分を投手に課すことができる。一方で、投手へのペナルティはデッドボールを与えた当該試合での退場のみであり、金銭罰則や次試合への出場停止といったものはない。このように投手はデッドボールをすることによって、打者に対して安全進塁権を与え、また場合によっては退場のリスクを背負う。

一方でなぜプロ野球の投手というピッチングの高度な技術を持った選手がデッドボールを投げてしまうのだろうか。その原因としては二つ考えられる。

【①アクシデント】

一つ目は打者の内角を攻める球は、打者に対して効果的な点である。内角を攻める球は打席のフォームを崩すという点において効果的であり、打者に近い軌道を描くため、コントロールが乱れるとアクシデント的に打者にあたってしまう。

【②戦略的デッドボール】

二つ目は、戦略としてのデッドボールが潜在的に許容されている点である。これは、先に述べたようにデッドボールによって打者は、怪我をして次打席以降出場できなくなり、また、その後のバッティングの調子を崩す可能性もある。そこで優秀な打者に対してデッドボールをあてることで、その打者の能力を抑制するものが戦略的デッドボールである。後者に関しては、モラルとして良くないことであり、現実的に存在しているのか疑問ではあるが、プロ野球選手の肌感覚としては存在しているようだ。例えば、メジャーリーグで活躍する上原浩治投手が大谷翔平の2018年のメジャー挑戦に対してあるコメントを残している。

「ただ、投手と野手の「二刀流」で挑戦する場合は『(相手チームが)デッドボールを当てにくる可能性もある』と指摘。『当てれば投手、野手、どちらかがつぶれる。メジャーでは、むしろ当てた方が褒められるかもしれない。覚悟はした方がいい』」 [3]

このように第一線球で活躍しているプロの投手でさえも、戦略としてのデッドボールの存在を許容している。しかし、これが野球のあるべき姿なのだろうか。150キロ近い硬球を選手に当てることを、是とすることが本来の野球ではないというのが今回の研究テーマを選んだ背景にある。

	正の影響	負の影響
投手	・相手打者のバッティングの調子を崩す可能性	・相手打者への安全進塁権 ・当該投手退場の可能性
打者	・安全進塁権	・選手生命にかかわるケガ ・バッティングの調子を崩す可能性

表 1.1 デッドボールによる投手・打者への影響

それでは年間プロ野球デッドボール数は一体どれくらいあるのだろうか。2016年ペナントリーグの年間デッドボール数は590個である。全体の試合数が両リーグ合わせて30年で80試合近く増加しているため、1試合あたりのデッドボール数で比較してみると、1990年代前半の1試合あたりのデッドボール数が0.5を切っているのに対して、2016年の1試合あたりのデッドボール数は0.7近くあることがわかる。このことから1試合あたりのデッドボール数はここ20年で微増していることがわかる。(図1)では、その原因とは何なのであろうか。上述したように、投手がデッドボールを投げる原因は、内角を攻める球の打者への効果と戦略的デッドボールの潜在的な許容にある。つまり、近年のデッドボール数の微増は、内角を攻める球の増加によるハプニング的なデッドボールの増加と、戦略的デッドボールの潜在的許容の広がりによるものと考えられる。

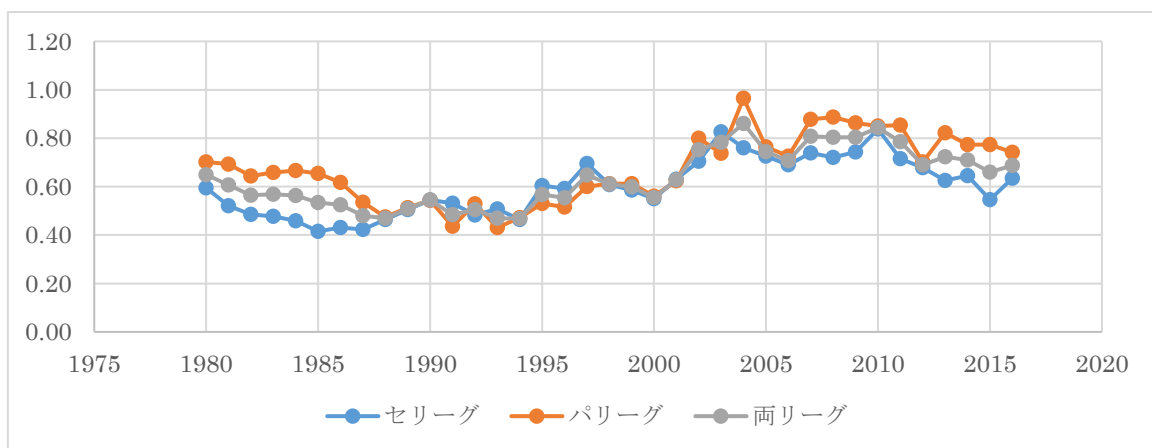


図1. 1 日本プロ野球1試合あたりのデッドボール数
出所) Baseball-Reference.com より筆者作成

ここまで、プロ野球選手が戦略的デッドボールを潜在的に許容していることを定性的に述べてきたが、定量的にも示唆できる。図 1.1 よりパリーグの方がセリーグより 1 試合あたりのデッドボール数が多いことが読み取ることができるが、実際に 1980 年～2016 年までのセリーグ・パリーグの 1 試合あたりのデッドボール数について T 検定をおこなうと P 値=0.02 となり「セリーグ・パリーグの 1 試合あたりのデッドボール数の母平均には差がある」ことがわかった。(図 1. 2) このような結果が出た理由としてパリーグの DH 制度が関係していると考えられる。つまり投手が打者としても打席に立つセリーグでは、投手のバッティング能力は低いため戦略的デッドボールを投げる必要性がない。一方 DH 制度を導入しているパリーグでは、投手は打席に立たず代わりに打撃専門の打者が一人分多いため、戦略的デッドボールの数が増加していると推察できる。

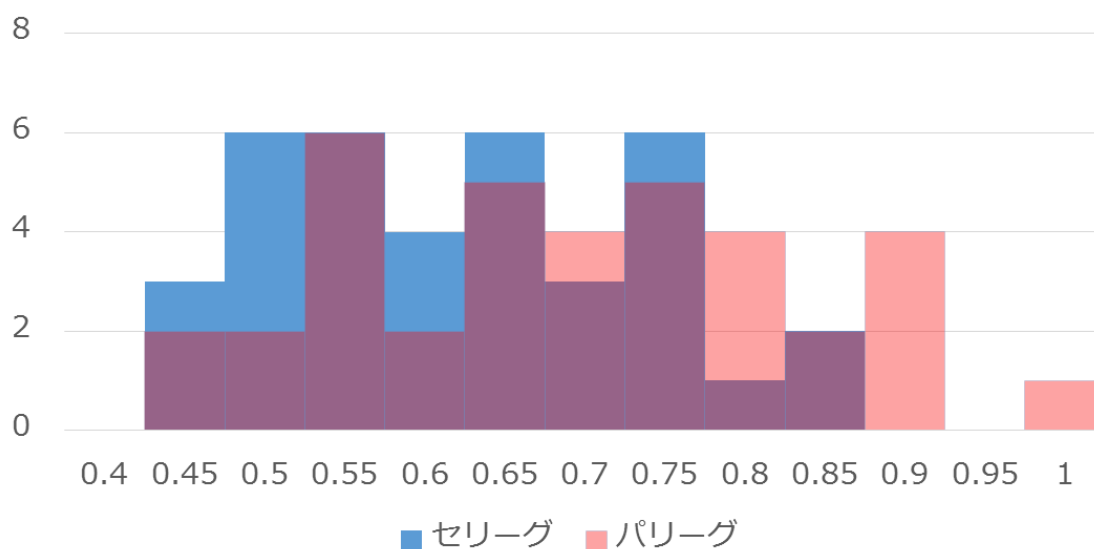


図 1.2 1980 年～2016 年までの 1 試合あたりの死球数分布

1.3 研究意義・目的

研究背景より、日本プロ野球が抱える課題として、「デッドボールは打者に対して負の影響があるにもかかわらず、日本プロ野球における 1 試合あたりのデッドボール数が増加していること」が浮かび上がってきた。そこで実際に打者に対してどのような負の影響があるのかを、選手のパフォーマンス指標で測定していく。また、選手全体のモラルの向上に繋げるペナルティの設計を行い、戦略的デッドボールの潜在的認知を無くしデッドボールの数を減らしていくことが、最終的な研究目的である。

2. 仮説

2.1 仮説設定

1.3 研究意義・目的より解決すべき課題を「デッドボールは打者に対して負の影響があるにもかかわらず、日本プロ野球における 1 試合あたりのデッドボール数が増加していること」と設定した。

では、上記の課題を解決するためには果たしてどのようにすればいいだろうか。まずデッドボールが打者のパフォーマンス指標に対して、どのような負の影響を与えるのかを測定する必要がある。次に、その測定した影響度を金銭的に評価してデッドボール 1 球の価値を算出することで、デッドボールに対するペナルティを設計し、戦略的デッドボールの潜在的認知を無くしデッドボールの減少という課題解決に繋げていく。(図 2.1)

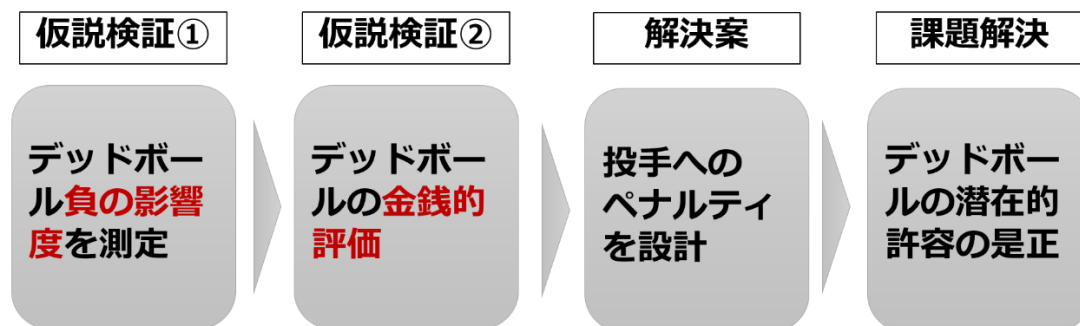


図 2.1 課題解決までのフロー

本研究では、この図 2.1 のフローに基づき 2 つの仮説検証を行う。

仮説①「デッドボールによって次打席以降の打者の出塁数は下がる」

デッドボール負の影響度を測定するにあたり、打者のパフォーマンス指標の設定を行う。打者のパフォーマンス指標は打率・出塁率・長打率・安打数・出塁数・打点など様々考えられるが、本研究では出塁数を分析対象とすることとする。その理由としては、出塁数は安打や打率といった打者の打力だけではなく、ボール球を見極め、四球を選択する選球眼の能力も含まれ、より広域な打者のパフォーマンス指標となり得ると考えられるからである。また、デッドボールによる影響を次打席だけではなく 10 打席まで測ることにした。これによりデッドボールによる打者への影響が何打席後まで継続するかを検証も行う。

仮説②「総出塁数の減少は翌年度選手年俵に負の影響を与える」

デッドボールの価値を算出するために出塁数が選手年俵に与える影響度の測定を行う。また、仮説①と関連付けるために総出塁数を変数に加えた賃金モデルを構築し係数の値から、デッドボールの金銭的に評価していく。

2.2 先行事例

セイバーメトリクスと名の付く研究は多く存在しているが、1.1で先述したように、野球が抱えている課題解決に統計手法を用いた研究は多く存在してはいない。そのため、本研究のようなデッドボールの影響度を測定し、それを金銭的に評価する一貫した先行研究はないが、デッドボールの発生確率に関しては「“It was a thought pitch” : Personal, situational, and target influences on hit-by-pitch events across time.」(2007、Timmerman, Thomas A.) [4]という研究がある。これは投手の視点から、相手チームがどのような攻撃パターンの時にデッドボールは発生するのかについて、その発生確率を比較している研究内容である。しかしながら、本研究は「打者視点である点」・「デッドボールの発生確率ではなくその後の影響度を測定している点」で大きく異なりこれまでにない新規性があると言える。

仮説②に関して、野球選手の賃金モデルを構築する先行研究については、野球選手の年俸を推定した「日米プロ野球投手の賃金関数の推定」(1993、樋口・山本) [5]という研究がある。

この研究は被説明変数を、自然対数をとった年俸として、説明変数を投手のパフォーマンス指標である「前年度防御率の逆数の自然対数値」・「前年度平均イニング数の自然対数値」、また選手の「勤続年数」・「リーグ順位の逆数」・「セリーグ・パリーグのダミー」・「FAダミー」とおいて、日米プロ野球投手の賃金関数モデルを構築している。この賃金モデルの説明変数の設定は大きく分けると、選手のパフォーマンス指標と勤続年数、所属チーム情報の3種類に分けることができる。本研究では、この先行研究をもとにその中のパフォーマンス指標と勤続年数を用いて、仮説②「総出塁数は翌年度選手年俸に正の影響を与える」ことを日本プロ野球選手の年俸モデルを構築していく。

2.3 データ収集

本研究のデータは、筆者が第6回スポーツデータ解析コンペティションに出場するにあたり、データスタジアム株式会社から提供を受けた。提供データの概要は、2014年・2015年のプロ野球ペナントレースの全打球・全打席・全試合ごとのデータである。今回は仮説①の検証のため全打席データのみを使用した。また、仮説②に使用する選手年俸と勤続年数に関しては、プロ野球選手データ名鑑[6]からデータの収集を行った。

3. 仮説①検証

3.1 データ概要

本章では、仮説①「デッドボールによって次打席以降の打者の出塁数は下がる」ことを検証していく。そこで2014年・2015年の日本プロ野球ペナントレースの全打席のデータを分析の対象とした。

データの処理として打席間隔が空くとデッドボールによる影響度が正確に測定できないことを鑑みて、2014年・2015年のペナントレースで上位124名の選手の打席データを抽出した。また、各選手の1打席目～10打席目までは出塁率が極端に高い場合と低い場合が見受けられたので、データから省くことにした。

今回の分析目的は、仮説である「デッドボールによって次打席以降の打者の出塁数は下がる」ことを示すことにある。そのため、デッドボールは何打席後まで打者の出塁の有無に影響を与えて、その影響はどれくらいであるのかの仮説検証をおこなう。今回仮説検証するにあたって設定した変数を以下に列挙する。

【被説明変数】

- ・ $Y(t)$ $t= 1,2,3,4,5,6,7,8,9,10$

ある年のある選手の1打席後～10打席後までの出塁の有無

選手が出塁した場合は $Y(t)= 1$ 、しなかった場合は $s(t)= 0$

【説明変数】

- ・ $X1$ $t=0$ 時点（当該打席）までのある年のある選手の出塁率

- ・ $X2$ $t=0$ 時点（当該打席）にデッドボールの有無

デッドボールを受けた場合は $X2= 1$ 、受けなかった場合は $X2= 0$

データ概要	日本プロ野球の全打席データ(131060行)
データ期間	2014年・2015年
データの種類	打席数上位124選手のクロスセクションデータ(66628行)
被説明変数 $Y(t)$	ある年のある選手の 1打席後～10打席後までの出塁の有無
説明変数 $X1$	ある選手のある年での当該打席までの出塁率
説明変数 $X2$	当該打席でのデッドボールの有無

表 3.1 データ概要のまとめ

変数	Y(t)	X1	X2
最小値	0.000	0.000	0.000
第1四分位	0.000	0.325	0.000
中央値	0.000	0.354	0.000
平均	0.345	0.354	0.009
第3四分位	1.000	0.382	0.000
最大値	1.000	0.833	1.000

表 3.2 変数の要約統計量

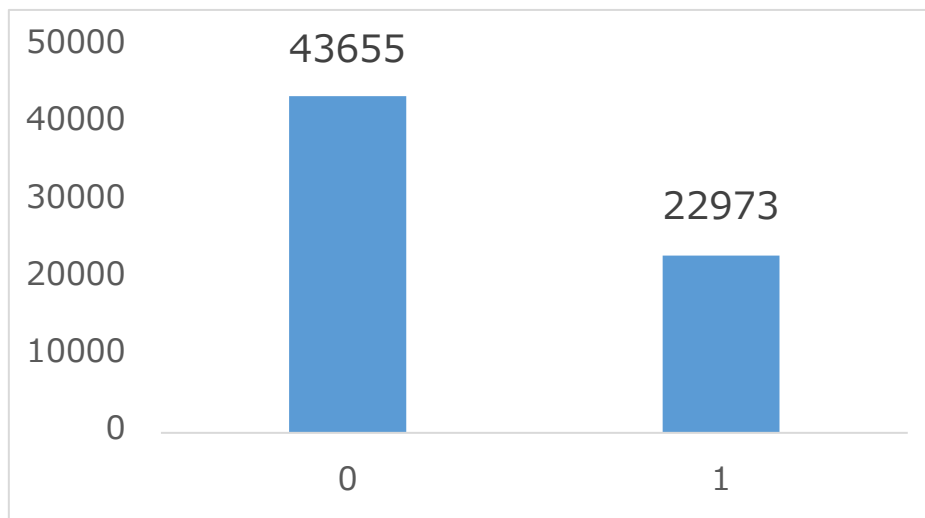


図 3.1 Y(t)のヒストグラム (横軸:デッドボールの有無、縦軸:頻度)

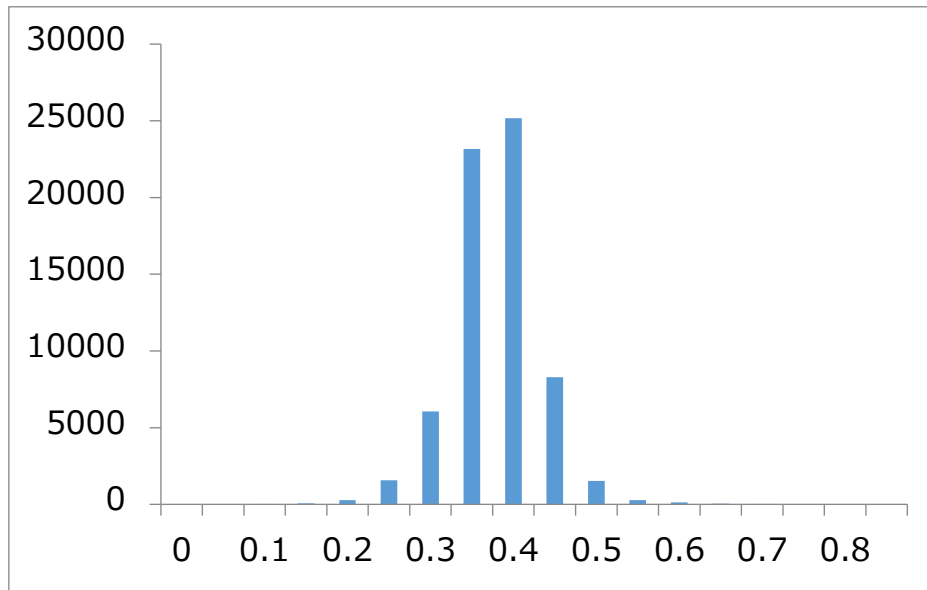


図 3.2 X1 のヒストグラム (横軸：分布、縦軸：頻度)

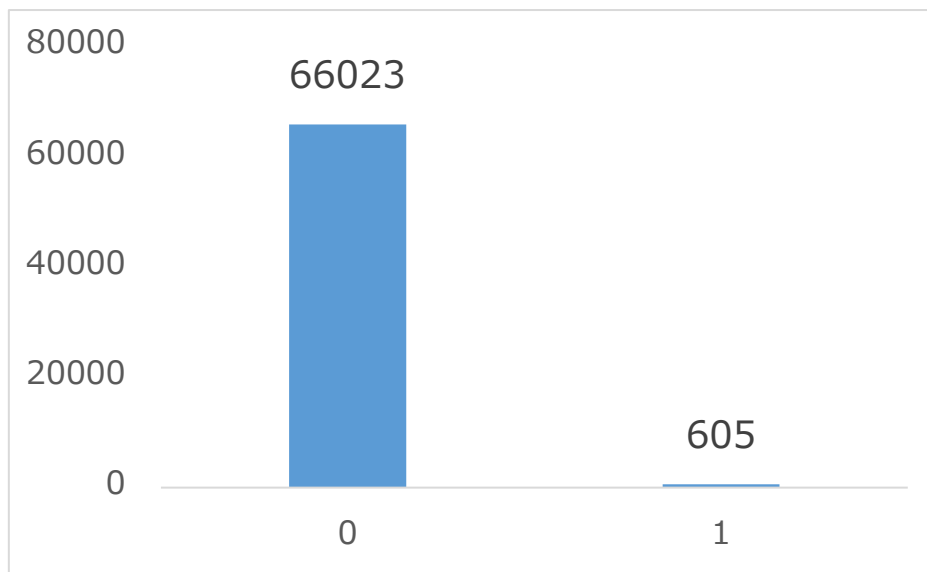


図 3.3 X2 のヒストグラム (横軸：分布、縦軸：頻度)

図 3.1 から選手の t 打席後出塁有無である被説明変数 $Y(t)$ がバイナリであること、図 3.3 から選手がデッドボールを受けたか否かの $X2$ がダミー変数であることがわかる。

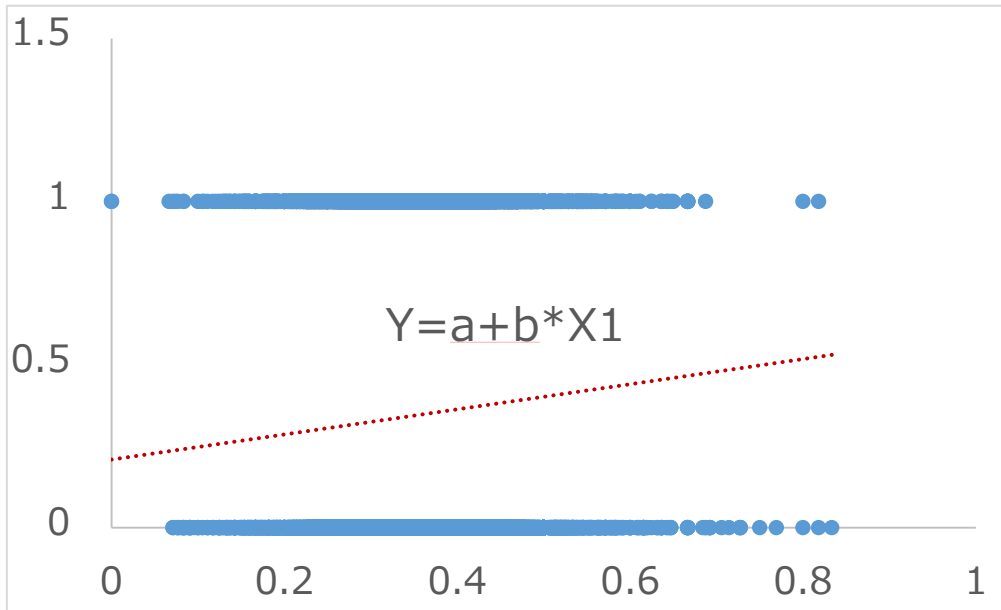


図 3.4 $Y(1)$ と $X1$ の散布図 (横軸 : $X1$ 、縦軸 : $Y(1)$)

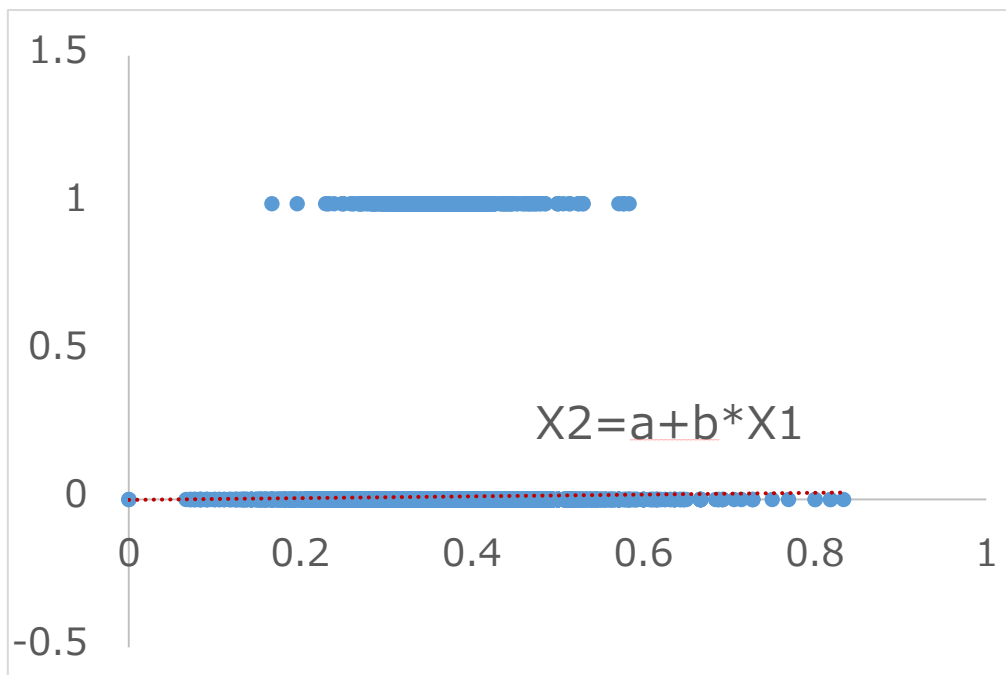


図 3.5 $X2$ と $X1$ の散布図 (横軸 : $X1$ 、縦軸 : $X2$)

図 3.4 より t 打席後の出塁有無と出塁率には相関がみられる。また、図 3.5 より $X1$ と $X2$ の説明変数間に相関は見られないことがわかる。

3.2 ロジスティック回帰

被説明変数が出塁したか（1）、しなかったか（0）のかのバイナリの変数であることからロジスティック回帰分析を使用して、結果のモデルからデッドボールが次打席期以降の出塁有無にどのように影響を与えているかの仮説検証を行う。ロジスティック回帰式は以下のように表すことができる。

【式】

$$P(Y(t) = 1 | X1, X2) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 X_1 + \beta_2 X_2)]}$$

(3.1)

以上の変数と式を用いてロジスティック回帰モデルを推定する。

被説明変数	Y(1)	Y(2)	Y(3)	Y(4)	Y(5)
β0	-1.247***	-1.231***	-1.226***	-1.226***	-1.241***
β1	1.709*** [0.385]	1.667*** [0.376]	1.645*** [0.371]	1.649*** [0.372]	1.688*** [0.381]
β2	-0.118 [-0.026]	-0.102 [-0.023]	0.147 [0.034]	-0.019 [-0.004]	0.054 [0.012]
AIC	85726	85743	85741	85744	85728
サンプルサイズ	66627	66627	66627	66627	66627

表 3.3 Y (t) =1~5 までのロジスティックモデル推定結果

注) ***印は 1%有意水準であることを示す。各変数の係数を掲載しており括弧内は限界効果

被説明変数	Y(6)	Y(7)	Y(8)	Y(9)	Y(10)
β0	-1.222***	-1.217***	-1.208***	-1.214***	-1.204***
β1	1.638*** [0.369]	1.621*** [0.366]	1.597*** [0.360]	1.613*** [0.364]	1.593*** [0.360]
β2	-0.010 [0.002]	0.154 [0.036]	0.011 [0.002]	0.039 [0.009]	-0.133 [-0.029]
AIC	85742	85743	85742	85741	85760
サンプルサイズ	66627	66627	66627	66627	66627

表 3.4 Y (t) =5~10 までのロジスティックモデル推定結果

注) ***印は 1%有意水準であることを示す。各変数の係数を掲載しており括弧内は限界効果

	Y(1)	Y(2)	Y(3)	Y(4)	Y(5)	Y(6)	Y(7)	Y(8)	Y(9)	Y(10)
$\beta 2$	-0.118	-0.102	0.147	-0.019	0.054	-0.010	0.154	0.011	0.039	-0.133
限界効果	-0.026	-0.023	0.034	-0.004	0.012	0.002	0.036	0.002	0.009	-0.029
P値	0.179	0.240	0.081	0.820	0.527	0.905	0.067	0.900	0.641	0.130

表 3.5 デッドボール係数 $\beta 2$ とその P 値

ロジスティックモデル推定結果から、出塁率は 1 打席後～10 打席後の打者の出塁有無に正の影響を与えることがわかる。(表 3.3, 表 3.4)

デッドボール係数である $\beta 2$ の値は、デッドボール後 2 打席目までの打者の出塁有無に負の影響を与えているが、3 打席後・5 打席後・7 打席後・8 打席後・9 打席後では正の影響を与えていることがわかる。また、デッドボール係数である $\beta 2$ の P 値は $s1 \sim s10$ すべての推定結果で 5%有意水準を満たしていない。(表 3.5)

【限界効果について】

仮説検証で得られたモデルから理論出塁確率を出してデッドボールの影響度を測定する。Y (1) の時のモデルについて出塁率の影響を除外した時のデッドボールによる出塁確率への影響度を Pd1 と設定する。

【式】

$$Pd1(X1) = P(Y(1) = 1 \mid X1, X2 = 1) - P(Y(1) = 1 \mid X1, X2 = 0)$$

(3.2)

限界効果とはこの Pd1 の平均のことである。つまり Y (1) のモデルでは、その他の変数条件が一定である場合、デッドボール変数 X1 の単位が 1 増加すると出塁確率が 2.6%下がることを示している。(表 3.3)

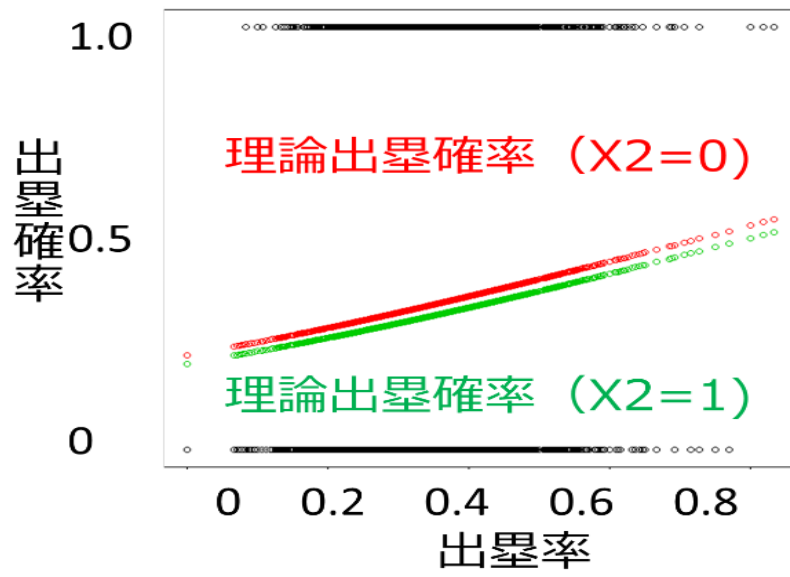


図 3.1 Y (1) のロジスティック回帰モデルを用いた理論出塁確率のグラフ

3.3 仮説①結果考察

3.2 ロジスティック回帰分析の結果より、デッドボールを受けた打者の1打席後の出塁確率は2.6%、2打席後の出塁確率は2.3%下がることがわかった。3打席後の係数 β_2 はプラスになっているが、3打席後ともなると試合をまたぐことで選手が復調するとも読み取れる。また、4打席以降はP値が高くデッドボールによる影響が希薄化しているため、デッドボールによる影響はないと判断した。このような結果になった理由として、二つ考えられる。

一つ目は、デッドボールが打者に当たる部位によって、身体的・精神的なダメージの負担が異なることが考えられる。具体的には頭部死球など打者にとって、深刻なダメージを与える場合、その後の打席にも負の影響を与えるが、ユニフォームにかすった場合は、打者にとってダメージはさほど大きくなく影響度は軽微なものである。デッドボール1球をとっても、このように当たる部位によって、打者に与える影響度は異なり、それらを同一視して分析したため係数 β_2 が有意な値にならなかったと推察できる。

二つ目として、デッドボールに対する打者のリアクションは統一でないことが考えられる。具体的にはデッドボールを受けたことによって、仮説の通りダメージの蓄積や恐怖心から積極的なバッティングができなくなる選手もいれば、対抗心を燃やして打ちに行く選手もいるということである。このように選手によって、デッドボールに対するリアクションは異なるため係数 β_2 が有意な値にならなかったと推察できる。

今回、分析結果から係数 β_2 に関して有意な値は得られなかったものの、少なくともデッドボールによる影響は2打席後までで、1打席後の出塁確率を2.6%、2打席後の出塁確率を2.3%下げると言えそうである。そこからデッドボール1球によって出塁数は0.049個下がるとして、仮説検証②に繋げていく。

4. 仮説②検証

4.1 データ概要

3.3 よりデッドボール 1 球によって打者の出塁数は 0.049 個下がることがわかった。しかしそれだけでは最終目標であるデッドボール 1 球の価値評価には不十分である。本章では、仮説②「総出塁数の減少は翌年度選手年俵に負の影響を与える」を検証し、出塁数と年俵の予測モデルからデッドボール 1 球の価値を算出し、ペナルティの設計につなげていく。分析の対象とするデータは 2014 年・2015 年のペナントレースの選手打席数上位 124 名の翌年度推定年俵、年間総出塁数、勤続年数のクロスセクションデータである。今回仮説検証するにあたって用いた変数を以下に列挙する。

【被説明変数】

- ・ $\ln Y$ ある年のある選手の翌年度の推定年俵の自然対数値

【説明変数】

- ・ $X1$ ある年のある選手の総出塁数
- ・ $X2$ ある年のある選手の勤続年数

データ概要	日本プロ野球選手の年間成績(124行)
データ期間	2014年・2015年
データの種類	打席数上位124選手のクロスセクションデータ
被説明変数 $\ln Y$	ある年のある選手の 翌年度推定年俵の自然対数値
説明変数 $X1$	ある選手のある年での総出塁数
説明変数 $X2$	ある選手のある年での勤続年数

表 4.1 データ概要のまとめ

変数	Y	X1	X2
最小値	1250	30	2.0
第1四分位	4800	164	7.0
中央値	8250	190	9.5
平均	13370	188	10.0
第3四分位	20000	215	13.3
最大値	60000	284	22.0

表 4.2 要約統計量

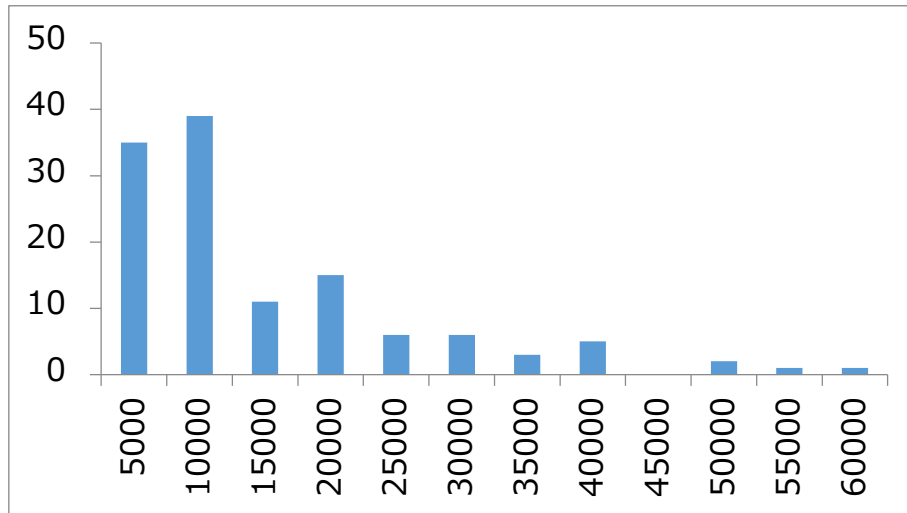


図 4.1 翌年度選手年俵 Y のヒストグラム (横軸：年俵、縦軸：頻度)

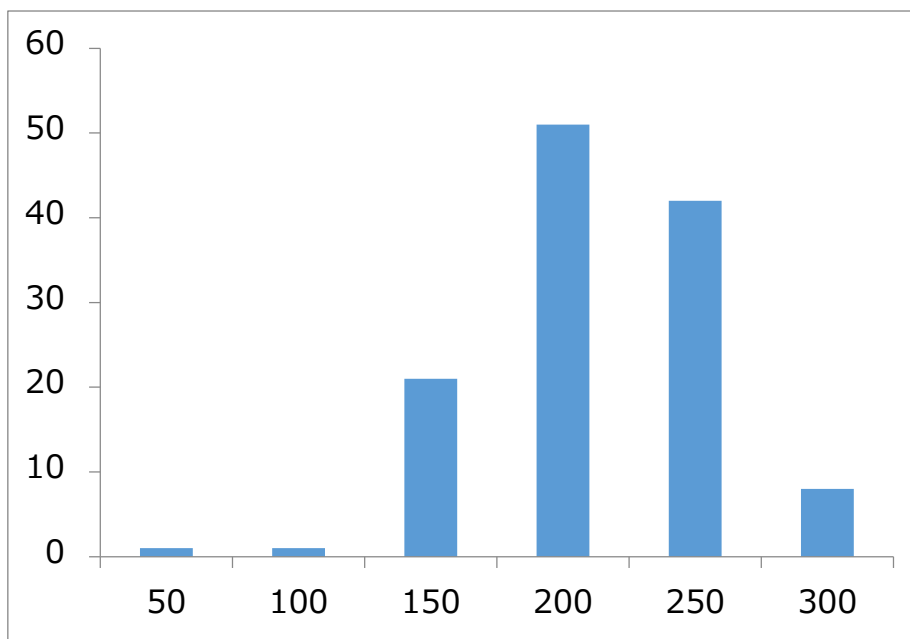


図 4.2 総出塁数 X_1 のヒストグラム (横軸：総出塁数、縦軸：頻度)

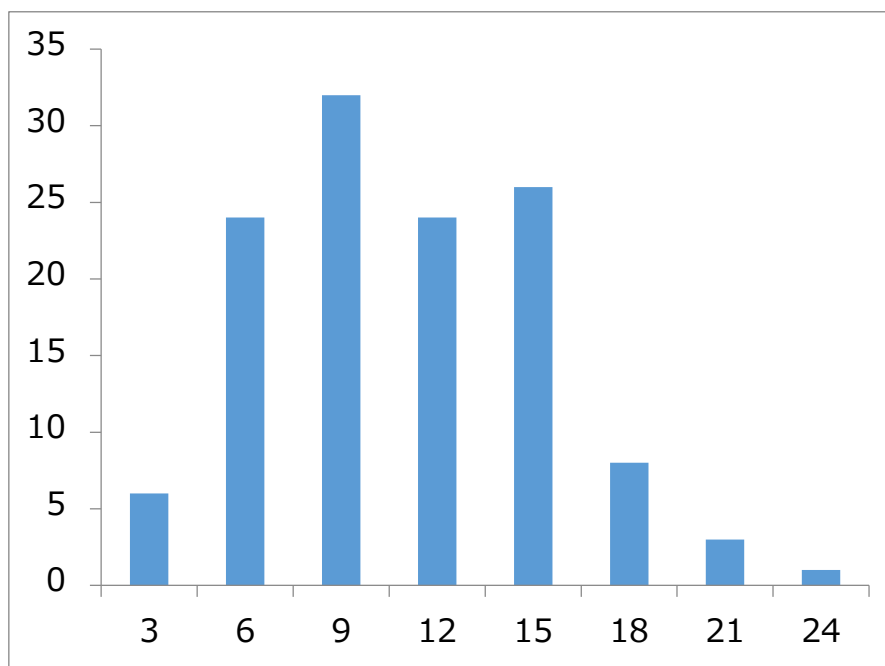


図 4.3 勤続年数 X2 のヒストグラム (横軸：勤続年数、縦軸：頻度)

表 4.2 と図 4.1 より選手の翌年度推定年俸は、1000 万円代の選手から最大 6 億円の選手まで桁の異なる金額の年俸が存在していることがわかる。そこで線形回帰式で推定を行うにあたり、翌年度推定年俸 Y は自然対数値をとることとする。

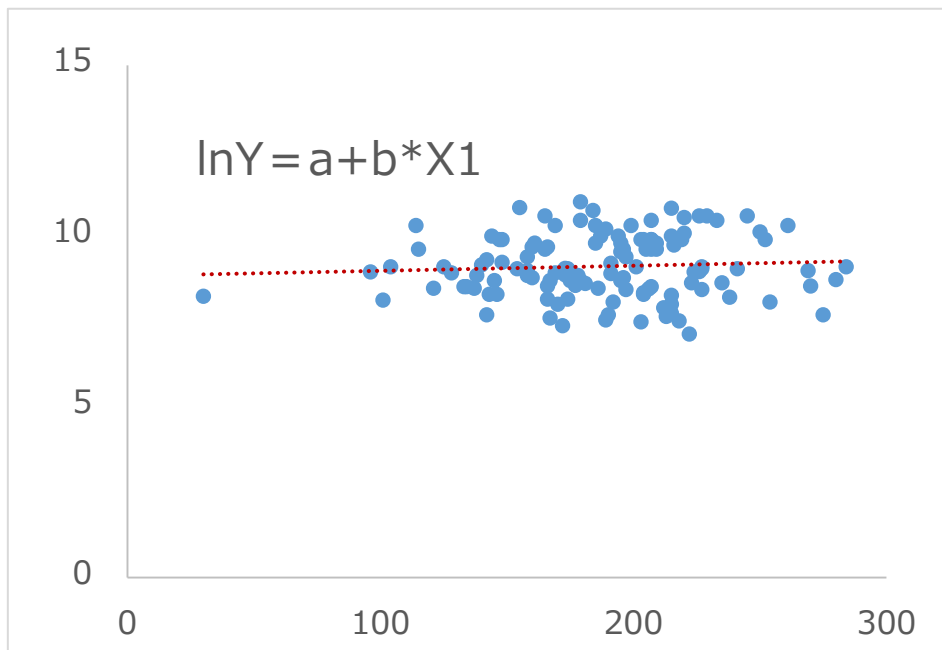


図 4.4 年俸（対数值）lnY と総出塁数 X1 の散布図
 （横軸：総出塁数 、縦軸：年俸（対数值））

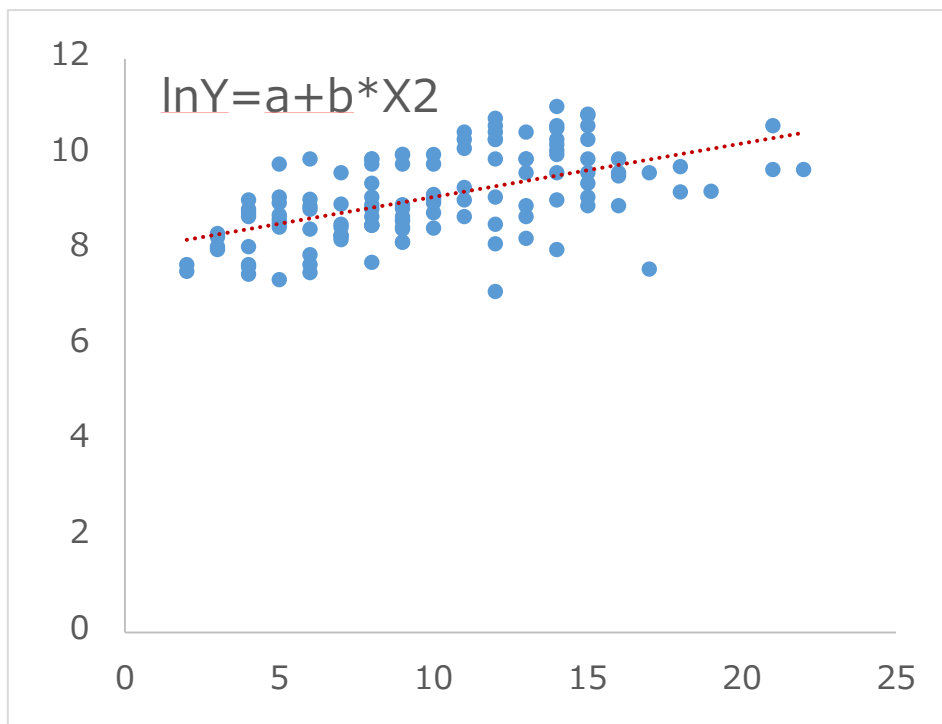


図 4.5 年俸（対数值）lnY と勤続年数 X2 の散布図
 （横軸：勤続年数 、縦軸：年俸（対数值））

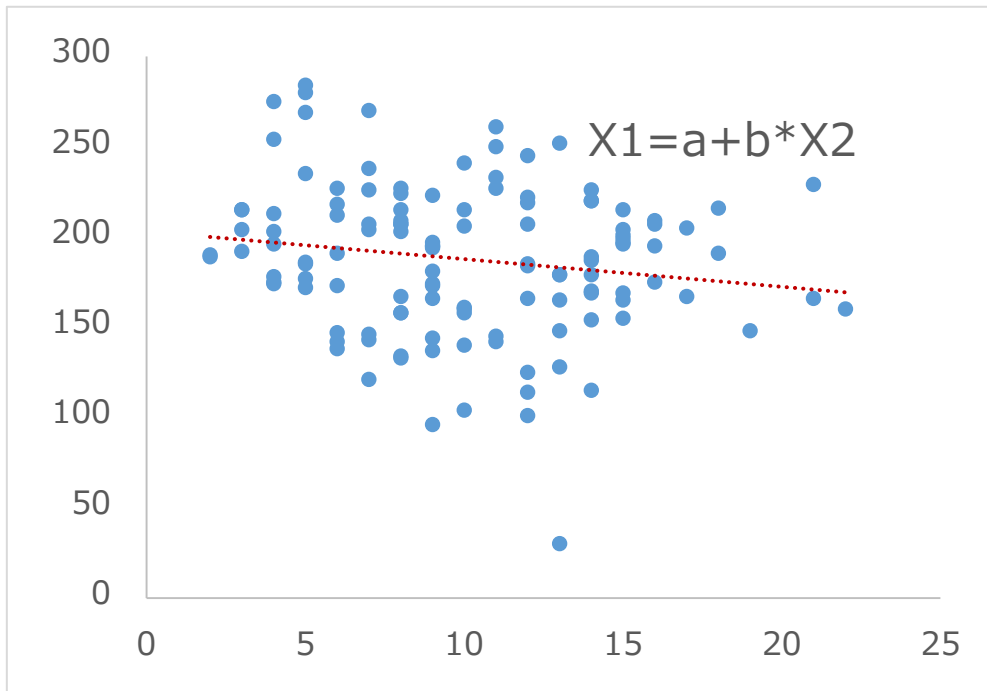


図 4.6 総出墨数 X1 と勤続年数 X2 の散布図
 (横軸：勤続年数 X2 、縦軸：総出墨数 X1)

図 4.4・図 4.5 より被説明変数 $\ln Y$ と説明変数 $X1 \cdot X2$ には正の相関関係があることが見られる。また、図 4.6 より説明変数間に相関は見られないことがわかる。

4.2 重回帰モデル

被説明変数を翌年度推定年俸としたことから重回帰分析を使用して、結果のモデルから総出塁数が翌年度年俸にどのように影響を与えているかの仮説検証を行う。重回帰式は以下のように表すことができる。

【式】

$$\ln Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

(4.1)

以上の変数と式を用いて重回帰モデルを推定する。

被説明変数＝選手年俸	翌年度選手年俸の自然対数値
切片	7.2711***
β_1 (前年度総出塁数)	0.0035**
β_2 (勤続年数)	0.1174***
自由度調整済決定係数	0.317
サンプルサイズ	124

表 4.3 重回帰モデル推定結果

注) ***, **印は 1%, 5% 有意水準であることを示す。(単位：万円)

表 4.3 より、この重回帰モデルは翌年度選手年俸の 31.7% を説明できるモデルである。また、他の変数条件が一定であるとき、総出塁数が 1 単位増えると翌年度推定年俸を 0.35% 増加させることがわかる。

4.3 仮説②結果考察

4.2 より、総出塁数が1回増えると翌年度推定年俸が0.35%増加することがわかった。また3.3より、デッドボール1球によって、打者の出塁数は0.049個減ることから、デッドボール1球によって打者が被る被害額（デッドボール1球の価値）は以下の式の通りとなる。

【式】

$$\begin{aligned}\text{デッドボール1球の価値} &= \text{年俸の} 0.35\% \times -0.049 \text{ 回} \\ &= -\text{年俸の} 0.017\%\end{aligned}$$

(4.2)

分析対象の平均である年俸1億3000万円の選手をモデルケースとすると、デッドボール1球の価値の以下の式の通りとなる。

【式】

$$\begin{aligned}\text{デッドボール1球の価値} &= 1 \text{ 億} 3000 \text{ 万円} \times - (0.017\%) \\ &= -2.2 \text{ 万円}\end{aligned}$$

(4.3)

(4.2) 式を用いてに、分析対象の選手の中で特にデッドボールによる被害額が大きかった上位5選手を表示すると（表4.4）のようになる。やはりクリーンナップを任されている優秀な選手は年俸が大きいことはもちろん、デッドボールを受ける回数も多い。ここからも戦略的デッドボールによる被害は優秀な打者に集中していることがわかる。

順位	2014	年俸	被害額	2015	年俸	被害額
1	ジョーンズ(12)	4億円	82万円	中島裕之(12)	4億円	82万円
2	阿部慎之助(7)	6億円	72万円	イデホ(9)	5億円	77万円
3	イデホ(10)	4億円	69万円	阿部慎之助(8)	5億1千万円	70万円
4	村田修一(13)	3億円	67万円	糸井嘉男(8)	3億5千万円	48万円
5	糸井嘉男(14)	2億5千万円	60万円	中村剛也(7)	3億8千万円	46万円

表 4.4 デッドボール被害額トップ5選手

(注) 括弧内の数字はその年に受けたデッドボールの数

5. 総括

5.1 考察のまとめ

3章仮説①検証と4章仮説②検証の結果をまとめる。

①デッドボールによって打者の1打席後の出塁確率を2.6%、2打席後の出塁確率を2.3%下げる。

②デッドボール1球によって打者の出塁数は0.049回減少する。

③他の変数条件が一定で、総出塁数が1回増えると、翌年度推定年俸は0.35%増加する。

④1億3000万円の選手のデッドボール1球による打者の被害額は-2.2万円である。

これらの結果を基に、各球団のデッドボールの年間被害額とプロ野球全体での年間被害額を算出すると（表5.1）のようになる。ここから日本プロ野球のデッドボールによる被害額は年間約400万円程であることがわかる。そこで、この数値を用いて選手のモラル向上を目的としたペナルティの設計を行う。

球団	2014			2015		
	平均年俸	死球数	被害額	平均年俸	死球数	被害額
巨人	6891	54	63	6893	52	61
中日	3538	34	20	3435	36	21
ヤクルト	2881	35	17	2926	37	18
阪神	3960	45	30	3558	38	23
広島	2679	50	23	3367	28	16
Dena	2462	61	26	2503	43	18
西武	3146	45	24	3314	64	36
ソフトバンク	5408	59	54	5798	51	50
日本ハム	3397	47	27	3259	46	25
ロッテ	3700	53	33	3269	44	24
オリックス	3295	85	48	4464	67	51
楽天	2923	45	22	2956	60	30
合計			388			375

表 5.1 球団別デッドボール被害額一覧

5.2 課題解決案

戦略的デッドボールの発生数の削減を目的として、デッドボールに対する次のようなペナルティ制度を導入することを提案したい。

【①ペナルティを支払うのは投手の所属球団】

まず、ペナルティの支払いに関して、投手自身に対してデッドボールのペナルティを課してしまうと、戦略的デッドボールは減少するが、それに伴って積極的な投球という強みまで奪ってしまう恐れがある。そこで、ペナルティを課す相手はデッドボールを当てた投手ではなく、投手の所属球団と設定する。

【②ペナルティの支払先はNPB（日本プロ野球機構）】

次に、ペナルティの支払先を設定する。デッドボールを受けた打者に対してペナルティを支払ってしまうと、打者がデッドボールに当たりに行くインセンティブが発生してしまうので、打者に直接払う仕組みは避けたい。もちろん、打者がデッドボールに当たりに行った場合、当該打者はルール上アウトになってしまうが、その見極めは非常に困難であるためそういった審判の負担を増やすことのないようにしなければならない。そこで、支払先はNPBと設定する。

【③NPB主催のデッドボールの危険性を伝える講演会・野球教室の運営】

最後にNPBに溜まった年間ペナルティ額およそ400万円の運用方法として、戦略的なデッドボールを非とするようなモラルのある次世代野球人の育成を提案する。具体的には、NPB主催のリトルリーグ・シニアリーグ（中学生野球）指導者に対するデッドボールの危険性を伝える講演会や、野球教室などを開催し、次世代の野球人のモラルの向上を図ることで、戦略的デッドボールの潜在的許容を減少させると考えられる。

6. 今後の課題

今回の分析結果は、筆者が当初予想していたデッドボールの価値よりかなり低い評価となった。また、筆者が本研究を発表したスポーツアナリティクス甲子園での審査員のコメントにも、過小評価している可能性が高いという意見もあった。今後、以下のことに留意して、より詳細な分析を行う必要がある。

【課題への対策】

- ・デッドボールが打者への影響度として出塁有無以外の変数も加える。

本研究では、打者のパフォーマンスとして出塁の有無のみをデッドボールが打者に与える影響度と設定したが、打者のパフォーマンスは出塁の有無だけではない。その他の守備や走塁といったパフォーマンスに対しても、デッドボールによる影響の大きさを測定することでより多面的に評価することができる。

- ・デッドボールによって故障した選手を分析対象に入れる。

本研究では、分析の対象者を打席の上位124名としている。したがって、デッドボールによって故障や調子を崩し試合に出場できなくなった選手に対しての影響度は測定できていない。そういった選手を分析の対象として加えることでより正確なデッドボールの価値を算出することができる。

【研究の発展】

- ・投手から見たデッドボールの価値評価

本研究では、分析対象を打者のみに限定していたが、デッドボールを与えてしまった投手に対してもその後のパフォーマンスに影響を与えている可能性がある。例えばデッドボールを投げてしまい、打者に怪我をさせた罪悪感から本来のピッチングができなくなるといったイップスを発症することもある。このようにデッドボールが受けた側だけではなく、与えた側に対してどのように影響を与えていくかを調べることで、新たな知見が得られる可能性がある。

- ・デッドボールが少ない年と多い年の比較

本研究では、デッドボールの原因として、①アクシデント②戦略的デッドボールの二つであると設定したが、デッドボールの数が少なかった年と多かった年の全打席データを比較することで、新たな原因を発見できる可能性がある。

7. 参考文献

- [1] 日本経済新聞 2013年9月28日朝刊37頁
- [2] 日本経済新聞 2013年9月18日朝刊27頁
- [3] デイリースポーツ 『上原浩治大谷へ「覚悟した方が」二刀流なら狙われる』
2016年12月14日
<http://www.daily.co.jp/mlb/2016/12/14/0009750099.shtml>
(アクセス日 2016年12月17日)
- [4] Timmerman, Thomas A. (著) 「“It was a thought pitch” : Personal, situational, and target influences on hit-by-pitch events across time.」、2007
- [5] 樋口美雄 (編集)・岡部浩二・岩瀬智彦・内田裕司・神谷之泰・河村宗寛・中村一成・山本勲 『プロ野球の経済学』, 日本評論社, 1993年
- [6] プロ野球選手データ名鑑 2012~2016, 別冊宝島社, 2012~2016年
- [7] 山本勲 (著), 「実証分析のための計量経済学」, 中央経済社, 2015年
- [8] 山田剛史・杉浦武俊・村井潤一郎 (共著), 「Rによるやさしい統計学」, オーム社, 2008年
- [9] 金明哲 (著), 「Rによるデータサイエンス」, 森北出版株式会社, 2007年
- [10] 豊澤栄治 (著) 「楽しいR-R 統計分析入門-」, 株式会社翔泳社, 2015年

8. 謝辞

本研究は情報・システム研究機構の新領域融合研究プロジェクト『社会コミュニケーション』データ中心科学リサーチコモンズ事業『人間・社会データ』の支援を受け作成することができました。また、データスタジアム株式会社からはデータの提供と修士論文へのデータ利用許諾をいただきました。

本研究を遂行するにあたり、林高樹教授からは手厚くご指導賜り、副査である河野宏和教授、林洋一郎准教授からもご専門に関連したご助言を賜りました。他にもゼミ同期生からの協力はもちろん、野球を愛する同学年の方々からも様々な応援・アドバイス頂きました。その結果として本研究は2016年スポーツアナリティクス甲子園パフォーマンス分析部門でSAP ジャパン賞を受賞することができました。この場を借りてご協力いただいた皆様に深く感謝申し上げます。

付表

本研究で用いた R のコードを記しておく。

```
new2 <- read_csv("C:/sopts_data/new2.csv",  
                na = "NA")
```

```
newdb <- na.omit(new2)
```

```
summary(newdb$s1)
```

```
summary(newdb$s2)
```

```
summary(newdb$s3)
```

```
summary(newdb$s4)
```

```
summary(newdb$s5)
```

```
summary(newdb$s6)
```

```
summary(newdb$s7)
```

```
summary(newdb$s8)
```

```
summary(newdb$s9)
```

```
summary(newdb$s10)
```

```
summary(newdb$syutu)
```

```
summary(newdb$dead_f)
```

```
d1 <- sum(newdb$dead_f)
```

```
d1
```

```
#ロジット回帰
```

```
dead.glm1 <- glm(y1~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm1)
```

```
dead.glm2 <- glm(y2~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm2)
```

```
dead.glm3 <- glm(y3~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm3)
```

```
dead.glm4 <- glm(y4~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm4)
```

```
dead.glm5<- glm(y5~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm5)
```

```
dead.glm6<- glm(y6~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm6)
```

```
dead.glm7<- glm(y7~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm7)
```

```
dead.glm8<- glm(y8~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm8)
```

```
dead.glm9<- glm(y9~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm9)
```

```
dead.glm10<- glm(y10~daritu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glm10)
```

```
#出墨率
```

```
dead.glms1<- glm(s1~syutu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glms1)
```

```
dead.glms2<- glm(s2~syutu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glms2)
```

```
dead.glms3<- glm(s3~syutu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glms3)
```

```
dead.glms4<- glm(s4~syutu+dead_f, family = binomial, data = newdb)
```

```
summary(dead.glms4)
```

```
dead.glms5<- glm(s5~syutu+dead_f, family = binomial, data = newdb)
```

```

summary(dead.glm5)

dead.glm6<- glm(s6~syutu+dead_f, family = binomial, data = newdb)
summary(dead.glm6)

dead.glm7<- glm(s7~syutu+dead_f, family = binomial, data = newdb)
summary(dead.glm7)

dead.glm8<- glm(s8~syutu+dead_f, family = binomial, data = newdb)
summary(dead.glm8)

dead.glm9<- glm(s9~syutu+dead_f, family = binomial, data = newdb)
summary(dead.glm9)

dead.glm10<- glm(s10~syutu+dead_f, family = binomial, data = newdb)
summary(dead.glm10)

#限界効果を算出
result1<-logitmfx(s1~syutu+dead_f, data=newdb, atmean=F)
result2<-logitmfx(s2~syutu+dead_f, data=newdb, atmean=F)
result3<-logitmfx(s3~syutu+dead_f, data=newdb, atmean=F)
result4<-logitmfx(s4~syutu+dead_f, data=newdb, atmean=F)
result5<-logitmfx(s5~syutu+dead_f, data=newdb, atmean=F)
result6<-logitmfx(s6~syutu+dead_f, data=newdb, atmean=F)
result7<-logitmfx(s7~syutu+dead_f, data=newdb, atmean=F)
result8<-logitmfx(s8~syutu+dead_f, data=newdb, atmean=F)
result9<-logitmfx(s9~syutu+dead_f, data=newdb, atmean=F)
result10<-logitmfx(s10~syutu+dead_f, data=newdb, atmean=F)

result1
result2
result3
result4
result5
result6

```

```
result7
result8
result9
result10
```

```
#チームを FACTOR に変換
```

```
newdb$bat_team_id<- factor(newdb$bat_team_id)
class(newdb$bat_team_id)
```

```
dead.glm4<- glm(andasu_f~b_daritu+b_dead+b_dead*bat_team_id, family =
binomial, data = newdb)
summary(dead.glm4)
```

```
plot(dead.glm4)
plot(newdb$b_daritu+newdb$b_dead, newdb$andasu_f)
```

```
#plot
plot(newdb$syutu, newdb$syutu_f, xlab="出墨率", ylab="出墨確率")
sy<- newdb$syutu
```

```
qq.0 <- predict(dead.glm4, data.frame( syutu = sy , dead_f = 0),
type="response")
qq.1 <- predict(dead.glm4, data.frame( syutu = sy , dead_f = 1),
type="response")
points(newdb$syutu, qq.0, col=2)
points(newdb$syutu, qq.1, col=3)
lines(newdb$syutu, qq.0, col=2)
lines(newdb$syutu, qq.1, col=3)
```

```
summary(qq.0 - qq.1)
```

```
t1<- which(newdb$dasekisu == 11)
t2<- t1-1
```

```
t3<- t2[-1]
newdb2<- newdb[t3,]
write.csv(newdb2, "newdb2.csv", quote=FALSE, row.names=FALSE, col.names
= TRUE)
```

```
summary(newdb2$salary)
summary(newdb2$syuturuisu)
summary(newdb2$among)
```

```
salary.lm<- lm(salary~syuturuisu+among, data = newdb2)
summary(salary.lm)
salary.lm2<- lm(salary~syuturuisu+deadsu, data = newdb2)
summary(salary.lm2)
```