

プロ野球本塁打数の分析

一般化線形モデル (ポアソン回帰モデル) の適用例

松尾 精彦

概要

一般化線形モデルとは、従来の線形回帰モデルを拡張したものである。具体的には、分布が正規モデルから、正規分布、ポアソン分布、二項分布、ガンマ分布などの指数型分布に拡張された点、連結関数 (*link function*) を導入することにより、平均パラメータが分布のパラメータの線形結合でなくても良くなった点、の2点である。この拡張により、これまで別々に扱われていたいくつかのモデルが同じプラットフォームで扱えるようになった。この論文では、一般化線形モデルの中でもよく用いられるポアソン回帰モデルの適用例を紹介する。データはチームの対戦チーム別本塁打数である。ポアソン回帰モデルがデータをよく説明できることから、副次的に“東京ドームでは風が巨人に有利に吹くか”についても検証する。

キーワード：一般化線形モデル，ポアソン回帰モデル，本塁打数，東京ドーム風問題
 経済学文献季報分類番号： 16 - 10

1 はじめに

一般化線形モデルは、正規線形モデルはもちろん、ロジスティックモデル、ポアソン回帰モデルといったモデルを、別々のものではなく、一つのプラットフォームで扱える範囲内で一般化したものである。この事実は統計解析ソフトウェアのコマンド (メニュー) に反映されている。Stata, R, S-Plus, Glim, Genstat, Systat, SAS-JMP といったソフトを用いる際には、一般化線形モデル (Generalized Linear Model) の命令に、オプションとして分布と結合関数とを与えると分析が行われる。この一般化線形モデルの利用は、特に分布が二項分布やポアソン分布に従うと仮定するとき、確率変数の正規変換をすることなく分析が可能になる。

第2節では、一般化線形モデルの定義を示し、ポアソン回帰モデルが一般化線形モデルに属していることを示す。次に、第3節ではホームランデータとそのモデル化を行う。最後

に、第4節では実際に分析を行った結果を示す。

2 一般化線形モデルの枠組みの中のポアソン回帰モデル

一般化線形モデルの定義は、[3]によると、独立な確率変数 Y_1, Y_2, \dots, Y_n が次の条件を満たすものである。

1. 各確率変数 Y_i の共通分布は正準形であり、下の式のような形をしている。

$$f(y_i; \theta_i) = \exp \left[\frac{y_i \theta_i - b_i(\theta_i)}{a_i(\phi)} + c_i(y_i, \phi) \right] \quad (1)$$

ここで、 $a_i(\phi) = \phi/n_i$ (n_i は既知定数) である。

2. $E[Y_i] = \mu_i$ とするとき、 $\mu_1, \mu_2, \dots, \mu_n$ は、 $g(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta}$ と表される。ここで $g(\cdot)$ は既知の狭義単調で微分可能な関数であり、連結関数 (link function) と呼ばれる。また $\boldsymbol{\beta}$ は p 次元未知パラメータベクトルで推定すべきもの、 \mathbf{x}_i は i 番目の確率変数に対応する p 次元説明変数ベクトルであり既知とする。なお、 \mathbf{x}_i^T は \mathbf{x}_i を転置の転置を表す。

上の定義は [1] の定義とは少し形が違う。その理由は分布の平均パラメータの構造には興味があるが局外パラメータ (例えば正規線形モデルにおける分散 σ^2) が存在するモデルも一般化線形モデルに含めるように工夫したものであり、また、[1] では正準形ではないモデルを含むように拡張されている。しかし、ここで扱うモデルではその違いを区別する必要はない。正準形をもつとは、 y と θ が、 $y\theta$ となっているモデルで、一般的には $a(y)\theta$ である。

本稿で分析に用いるモデルはポアソン回帰モデルであり、ポアソン分布と対数リンク関数を持つモデルである。すなわち、

$$f(y_i; \mu_i) = \frac{\mu_i^{y_i}}{y_i!} e^{-\mu_i}, \quad \log(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta} \quad (2)$$

上の定義の左部分は、

$$f(y_i; \mu_i) = \exp\{y_i \log \mu_i - \mu_i - \log y_i!\} \quad (3)$$

と表され、(1) を満たすことが分かる。

3 データとモデル化

前節において、ポアソン回帰モデルが一般化線形モデルの構造を持つことを示した。この説ではプロ野球セリーグの、2008年から2010年のシーズンにおける、ホームラン数がポアソン回帰モデルに従っていることを示す。本稿の最後に配した付表1～3は、[5]～[7]に記載されている各試合の記録から、各チームが各球場で各対戦チームと戦ったときのホームラン数をまとめたものである。このデータの応答変量は本塁打数であり、それを説明するのが、試合数、チームの打力、対戦チームの投手力、そして球場の形状と広さと考える。つまり a_i をチームの攻撃力、 b_j を対戦チームの投手力、 c_k を球場の効果とし、 n_{ijk} をこの条件で戦った試合数としよう。すると、ホームラン数の期待値は、

$$\mu_{ijk} = n_{ijk} \exp\{a_i + b_j + c_k\} = \exp\{\log(n_{ijk}) + a_i + b_j + c_k\} \quad (4)$$

と表される。なお、 a_i, b_j, c_k は質的な要因であり、どれか一つの水準が0になるように制約を与えている。また、 $\log(n_{ijk})$ はオフセットでありパラメータは伴わない。

さて、上記の期待値を持つホームラン数がポアソン分布に従うという仮定は、ホームランの性質から見て妥当である。なぜならばホームランの出る確率は低く、1試合毎に数十回の打席が各チームに与えられることを考えれば、希現象の大量観察としてポアソン分布を想定する根拠があるからである。いずれにせよ、この仮定は分析を通して明らかにされる。

4 分析と結果

前節のモデルが実際のデータに適合するかを、2008、2009、2010の3年間について調べてみる。データはセ・リーグチーム同士が、どちらかのホームグラウンドで戦ったものだけを集めたものであり、長野球場、福井フェニックス球場といった少数の試合しか行われなかった球場での試合や、セ・パ交流戦を除外している。なお、ここで行った分析はすべてSAS JMP Pro 11を用いて得られたものである。

4.1 2008年度シーズン

まず2008年シーズンのモデルのあてはまり具合を見ると、

モデル全体の検定				
モデル	(-1)* 対数尤度	尤度比カイ 2 乗	自由度	p 値 (Prob>ChiSq)
差分	57.1311231	114.2622	15	< .0001
完全	138.641636			
縮小	195.772759			

であり、チームの打力、対戦チームの投手力、球場の効果を加えたモデルは、全体としてデータをよく説明していることが分かる。なお、上の表の完全というのは、すべての要因を含むモデルにおける最大対数尤度にマイナス 1 をかけた量であり、縮小とは切片のみを用いたモデルにおける最大対数尤度にマイナス 1 をかけた量である。これらの差を計算したものが差分であり、その差分を 2 倍したものが尤度比カイ 2 乗である。この量は漸近的にカイ 2 分布にしたがうことが分かっている、すべての要因を加えたモデルが有意であるかどうかの p 値を与える。この場合、自由度 15 にたいし、漸近的にカイ 2 乗分布に従う尤度比カイ 2 乗統計量の値が 114.2622 であり、その p 値は < .0001 なので、すべての要因を含むモデルが高度に有意であることが示される。

続いて、モデルの適合度をみてみよう。ポアソン回帰モデルでは分布がポアソン分布であるため、局外パラメータ (nuisance parameter) が存在しない。そのため適合度検定が可能となり、以下の表のようにまとめられる。

適合度統計量	カイ 2 乗	自由度	p 値 (Prob>ChiSq)
Pearson	41.4151	44	0.5830
デビアンス	42.0900	44	0.5538

上の表は、飽和モデルとの比較を行い、モデルが全変動をうまく説明しているかを示すものである。 p 値が小さいとき、モデルの適合度が悪くなる。この年度では、適合度統計量を Pearson としてもデビアンスとしても、ポアソン回帰モデルがデータをうまく説明していることが分かる。

次に各要因の効果の検定表をみてみよう。下の表でもわかるように、チームの打力、対戦チームの投手力、球場の効果がいずれも有意であることが分かる。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	20.86639	0.0009
対戦チーム	5	15.065468	0.0101
球場	5	35.92213	<.0001

以上を総合すると、各チームの、対戦チーム別球場別本塁打数はポアソン分布に従っていることが分かる。そこで、各要因の各水準のパラメータ推定値をもとに、どの球団の打撃力が強いのか、どの球団の投手力が強いのか、どの球場で本塁打が出やすいのかについて見てみよう。

パラメータ推定値				
項	推定値	標準誤差	尤度比カイ 2 乗	p 値
切片	-0.271188	0.0482238	35.620111	<.0001*
チーム [ヤクルト]	-0.28284	0.1245034	5.4982858	0.0190*
チーム [横浜]	0.1638741	0.1137017	2.0137551	0.1559
チーム [巨人]	0.2194265	0.1072313	4.0043477	0.0454*
チーム [広島]	-0.281103	0.1287822	5.082561	0.0242*
チーム [阪神]	-0.116653	0.1218942	0.9446334	0.3311
対戦チーム [ヤクルト]	0.1447135	0.1065721	1.7878925	0.1812
対戦チーム [横浜]	0.3232797	0.1079049	8.4335885	0.0037*
対戦チーム [巨人]	-0.206048	0.1221945	2.9626754	0.0852
対戦チーム [広島]	-0.153032	0.1235154	1.590733	0.2072
対戦チーム [阪神]	-0.19843	0.1239389	2.7038004	0.1001
球場 [ナゴヤドーム]	-0.215083	0.1215473	3.1549567	0.0757
球場 [横浜スタジアム]	0.1906099	0.1136524	2.8362301	0.0922
球場 [広島市民]	0.2318705	0.1328136	3.0787431	0.0793
球場 [甲子園]	-0.704195	0.1702598	19.016262	<.0001*
球場 [神宮]	0.0303863	0.1246774	0.0593715	0.8075

上の表から、巨人と横浜の打力が最も強力である一方でヤクルトと広島が打力が劣っていることがわかる。投手力をみると、巨人、広島そして阪神の投手力が充実している一方で横浜とヤクルトの投手力が劣っていることが分かる。最後に球場を見ると甲子園球場とナゴヤドームでは本塁打が出にくいことが分かる。

ついでに、巨人が東京ドームで、全体の傾向に輪をかけて、本塁打が多いのかどうかを調べてみる。巨人が東京ドームで戦う場合に説明変数 1 を与えるダミー変数をモデルに加えた場合の効果の検定は下の表のようになる。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	20.154486	0.0012
対戦チーム	5	14.137123	0.0148
球場	5	27.378372	< .0001
巨人ドーム	1	0.0001094	0.9917

上の表で、巨人ドームの効果の p 値が 0.9917 であることから、東京ドームで巨人に有利な風が吹いているとは言えない。

4.2 2009 年度シーズン

2009 年度シーズンでは、広島のフランチャイズ球場がマツダスタジアム広島に変わり、前年度とは球場の効果に変化がみられることが予想される。2008 年度と同様に、チームの打力、対戦チームの投手力、そして球場の効果を調べてみよう。最初に、チーム力、対戦チーム、球場のすべての要因を含むモデルの適合度を調べることにしよう。まず、チーム、対戦チーム、球場の 3 つの要因を含むモデルの効果は、

モデル全体の検定				
モデル	(-1)* 対数尤度	尤度比カイ 2 乗	自由度	p 値 (Prob>ChiSq)
差分	26.7319465	53.4639	15	< .0001
完全	149.022887			
縮小	175.754833			

は高度に有意であり，要因すべてを含むモデルは妥当であり，また，モデルの適合度は，

適合度統計量	カイ 2 乗	自由度	p 値 (Prob>ChiSq)
Pearson	56.3509	44	0.1003
デビアンズ	56.1402	44	0.1037

と， p 値が 0.1 をわずかに上回り，モデルのあてはまりも悪いとは言えない．次に各要因の水準の効果を見てみよう．

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	13.3627	0.0202*
対戦チーム	5	5.0898068	0.4050
球場	5	20.699136	0.0009*

前年とは違って，対戦チームの投手力が有意でなくなっている．そこで 3 つの要因すべてを含むモデルと，対戦チームを除いたモデルの 2 つのモデルを考える必要があるかもしれない．まず，3 つの要因をすべて含むモデルでは，

パラメータ推定値				
項	推定値	標準誤差	尤度比カイ 2 乗	p 値
切片	-0.161559	0.0431393	14.868957	0.0001*
チーム [ヤクルト]	-0.152557	0.1237215	1.5729865	0.2098
チーム [横浜]	-0.116446	0.1155062	1.0414858	0.3075
チーム [巨人]	0.2568468	0.1027456	5.9420222	0.0148*
チーム [広島]	-0.067426	0.117239	0.3359565	0.5622
チーム [阪神]	-0.156552	0.116814	1.8639779	0.1722
対戦チーム [ヤクルト]	0.0145779	0.1173271	0.0153869	0.9013
対戦チーム [横浜]	0.0887364	0.1077633	0.6654946	0.4146
対戦チーム [巨人]	-0.01736	0.1122761	0.0239881	0.8769
対戦チーム [広島]	0.0961509	0.1105708	0.7398369	0.3897
対戦チーム [阪神]	-0.244578	0.1200806	4.3984058	0.0360*
球場 [ナゴヤドーム]	-0.371364	0.1216124	9.5460656	0.0020*
球場 [マツダスタジアム広島]	-0.271307	0.1271769	4.5841524	0.0323*
球場 [横浜スタジアム]	0.1888579	0.11562	2.6786833	0.1017
球場 [甲子園]	0.0268207	0.1323647	0.0410214	0.8395
球場 [神宮]	0.1599249	0.1231781	1.7044793	0.1917

巨人の打力が突出して高いことが分かる。また投手力では阪神の効果だけが有意になっている。球場では、ナゴヤドームとマツダスタジアム広島で本塁打が出にくくなっていることが分かる。以上のように、3つの要因すべてを含むモデルは、データによくあてはまっていると言える。

次に、対戦チーム要因を除いたモデルをあてはめてみる。

モデル全体の検定				
モデル	(-1)* 対数尤度	尤度比カイ 2 乗	自由度	p 値 (Prob>ChiSq)
差分	24.1870431	48.3741	10	<.0001
完全	151.56779			
縮小	175.754833			

とモデルは全体として高度に有意になっている。また、適合度も、

適合度統計量	カイ 2 乗	自由度	p 値 (Prob>ChiSq)
Pearson	62.4380	49	0.0941
デビアンズ	61.2300	49	0.1128

と適合度の p 値は、やや小さいのだが、まずまずの適合度を示している。効果の検定も下の表のようにチーム要因、球場要因ともに有意である。以上のことから、チーム要因、球場要因を持つポアソン回帰モデルはデータにあてはまっていることがわかる。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	16.130282	0.0065
球場	5	26.815988	< .0001

パラメータの推定値を見てみると、

パラメータ推定値				
項	推定値	標準誤差	尤度比カイ 2 乗	p 値
切片	-0.158456	0.0429491	14.408667	0.0001*
チーム [ヤクルト]	-0.157436	0.1075749	2.2087597	0.1372
チーム [横浜]	-0.159904	0.1031517	2.4818659	0.1152
チーム [巨人]	0.2649237	0.0900026	8.2954484	0.0040*
チーム [広島]	-0.099587	0.1091915	0.8509496	0.3563
チーム [阪神]	-0.067401	0.1055243	0.4143587	0.5198
球場 [ナゴヤドーム]	-0.342473	0.1098028	10.475164	0.0012*
球場 [マツダスタジアム広島]	-0.211454	0.1098465	3.8879695	0.0486*
球場 [横浜スタジアム]	0.2432204	0.0949733	6.2548333	0.0124*
球場 [甲子園]	-0.112295	0.1113543	1.0441569	0.3069
球場 [神宮]	0.1665797	0.0960413	2.9140382	0.0878

と打力では巨人が強力であり，球場ではナゴヤドーム，マツダスタジアム広島，甲子園が本塁打の出にくい球場で，横浜スタジアムが本塁打の出やすい球場であることが分かる。

この文節の最後に，巨人ドーム風効果を2つのモデルで観察しよう。まず，すべての要因を盛り込んだモデルでは，巨人ドーム風の効果は有意ではない。それどころか，推定値が -0.267855 とマイナスの値を取っているのである。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	14.368272	0.0134
対戦チーム	5	5.5543561	0.3520
球場	5	20.387175	0.0011
巨人ドーム	1	1.0115976	0.3145

次に対戦チーム要因を取り去ったモデルについて見てみよう。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	15.550676	0.0083
球場	5	26.133651	< .0001
巨人ドーム	1	0.5470483	0.4595

上の表でわかるように，このモデルにおいても巨人ドーム効果は有意ではない。パラメータの推定値も -0.16457 であり，東京ドームで巨人に有利な風が吹いているとはいえない。

4.3 2010 年度シーズン

前節，前々節と同様に，各チームの本塁打数をチーム，対戦チーム，球場の3つの要素を持つポアソン回帰モデルに当てはめ，そのあてはまりが良好であることを示す。この事実を副次的に利用して，東京ドームで巨人に有利な風が吹いているかどうかを検定する。

モデル全体の検定				
モデル	(-1)* 対数尤度	尤度比カイ 2 乗	自由度	p 値 (Prob>ChiSq)
差分	39.0825284	78.1651	15	< .0001
完全	142.830669			
縮小	181.913197			

であり、全体としてモデルは有効であり、また、適合度は、

適合度統計量	カイ 2 乗	自由度	p 値 (Prob>ChiSq)
Pearson	33.3566	44	0.8788
デビアンズ	34.2399	44	0.8547

と良好である。次に各要因の効果をみる。前年と同じく対戦チーム要因が有意ではなくなっている。しかも p 値が 1 に近い値になっている。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム名	5	31.99472	< .0001
対戦チーム	5	1.5086835	0.9121
球場名	5	22.769681	0.0004

そこで、2010 年度シーズンの分析は、対戦チーム要因を取り除くことにする。

モデル全体の検定				
モデル	(-1)* 対数尤度	尤度比カイ 2 乗	自由度	p 値 (Prob>ChiSq)
差分	38.3281866	76.6564	10	< .0001
完全	143.585011			
縮小	181.913197			

であり、全体としてモデルは有効であり、また、適合度は、

適合度統計量	カイ 2 乗	自由度	p 値 (Prob>ChiSq)
Pearson	35.2319	49	0.9304
デビアンズ	35.7485	49	0.9213

であり、ポアソン回帰モデルのあてはまりはかなり良好である。下の表を見て分かるように、チーム要因を取り去った、チーム・球場要因は高度に有意である。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム	5	38.5952	< .0001
球場	5	30.287835	< .0001

下の表を見ればわかるように打力は巨人が強力で、阪神がそれに続く。球場はナゴヤドームと甲子園では出にくくなっていて、マツダスタジアム広島では出やすくなっている。

パラメータ推定値				
項	推定値	標準誤差	尤度比カイ 2 乗	p 値
切片	-0.015535	0.0406975	0.1465327	0.7019
チーム [ヤクルト]	-0.109662	0.0987364	1.2612642	0.2614
チーム [横浜]	-0.139355	0.0998771	2.0031923	0.1570
チーム [巨人]	0.4121207	0.0818991	23.746155	< .0001*
チーム [広島]	-0.391585	0.106044	14.886105	0.0001*
チーム [阪神]	0.2015053	0.0901138	4.7994976	0.0285*
球場 [ナゴヤドーム]	-0.523042	0.1122124	24.76594	< .0001*
球場 [マツダスタジアム広島]	0.1840451	0.0902969	3.9964812	0.0456*
球場 [横浜スタジアム]	0.151177	0.0924679	2.592768	0.1074
球場 [甲子園]	-0.10402	0.0996656	1.1135976	0.2913
球場 [神宮]	0.1466411	0.0922105	2.4561948	0.1171

最後に、このモデルに基づき、巨人の東京ドーム効果を調べたが、下の表にあるように、有意ではない。しかし、推定値は 0.24486658 と正の値を取っており、本稿で扱った 3 年の

うちでは最も疑わしい年と言える。

効果の検定			
要因	自由度	尤度比カイ 2 乗	p 値 (Prob>ChiSq)
チーム名	5	26.509915	< .0001
球場名	5	28.213997	< .0001
巨人ドーム	1	1.3833819	0.2395

5 おわりに

当初は問題ありきで、巨人が東京ドーム風を利用してホームランを量産しているという噂を統計的に検定できないかという動機付けがあった。その答えを得るには各チームの本塁打数にモデルをあてはめなければならない。ホームラン数のデータを眺めていて、一試合当たりのホームラン数はポアソン分布に従っていること、一試合当たりのホームラン数の平均は、チームの打力、対戦チームの投手力、試合の行われる球場の3つの要因で説明できるのではないかと考えた。この仮定を実証するためにデータを収集し、ポアソン回帰モデルがデータをうまく説明しているかを確認したところ、うまく適合することが分かった。

このようなモデルのあてはめは、様々な仮説を検証する土台となる。つまり説明変数やパラメータを追加することの妥当性を検定できるということだ。本稿では、東京ドームの風は巨人に味方をするかという問題に対し、「味方をする」という仮説を採択する積極的な理由はないとの答え得た。

付表 1. 2008 年度ホームラン数

チーム	対戦チーム	球場	試合数	合計ホームラン数	オフセット
ヤクルト	横浜	横浜スタジアム	12	15	2.485
ヤクルト	横浜	神宮	12	10	2.485
ヤクルト	巨人	神宮	12	4	2.485
ヤクルト	巨人	東京ドーム	12	12	2.485
ヤクルト	広島	広島市民	11	6	2.398
ヤクルト	広島	神宮	12	4	2.485
ヤクルト	阪神	甲子園	11	3	2.398
ヤクルト	阪神	神宮	10	3	2.303
ヤクルト	中日	ナゴヤドーム	12	3	2.485
ヤクルト	中日	神宮	10	9	2.303
横浜	ヤクルト	横浜スタジアム	12	13	2.485
横浜	ヤクルト	神宮	12	18	2.485
横浜	巨人	横浜スタジアム	11	9	2.398

横浜	巨人	東京ドーム	9	6	2.197
横浜	広島	横浜スタジアム	10	10	2.303
横浜	広島	広島市民	12	13	2.485
横浜	阪神	横浜スタジアム	12	9	2.485
横浜	阪神	甲子園	6	3	1.792
横浜	中日	ナゴヤドーム	10	5	2.303
横浜	中日	横浜スタジアム	11	17	2.398
巨人	ヤクルト	神宮	12	15	2.485
巨人	ヤクルト	東京ドーム	12	26	2.485
巨人	横浜	横浜スタジアム	11	18	2.398
巨人	横浜	東京ドーム	9	19	2.197
巨人	広島	広島市民	12	12	2.485
巨人	広島	東京ドーム	9	11	2.197
巨人	阪神	甲子園	12	5	2.485
巨人	阪神	東京ドーム	12	17	2.485
巨人	中日	ナゴヤドーム	12	8	2.485
巨人	中日	東京ドーム	10	10	2.303
広島	ヤクルト	広島市民	11	6	2.398
広島	ヤクルト	神宮	12	5	2.485
広島	横浜	横浜スタジアム	10	12	2.303
広島	横浜	広島市民	12	11	2.485
広島	巨人	広島市民	12	10	2.485
広島	巨人	東京ドーム	9	10	2.197
広島	阪神	広島市民	9	4	2.197
広島	阪神	甲子園	9	1	2.197
広島	中日	ナゴヤドーム	12	3	2.485
広島	中日	広島市民	12	14	2.485
阪神	ヤクルト	甲子園	11	3	2.398
阪神	ヤクルト	神宮	10	6	2.303
阪神	横浜	横浜スタジアム	12	9	2.485
阪神	横浜	甲子園	6	3	1.792
阪神	巨人	甲子園	12	2	2.485
阪神	巨人	東京ドーム	12	11	2.485
阪神	広島	広島市民	9	8	2.197
阪神	広島	甲子園	9	3	2.197
阪神	中日	ナゴヤドーム	12	12	2.485
阪神	中日	甲子園	11	6	2.398
中日	ヤクルト	ナゴヤドーム	12	10	2.485
中日	ヤクルト	神宮	10	14	2.303
中日	横浜	ナゴヤドーム	10	12	2.303
中日	横浜	横浜スタジアム	11	17	2.398
中日	巨人	ナゴヤドーム	12	9	2.485
中日	巨人	東京ドーム	10	11	2.303
中日	広島	ナゴヤドーム	12	12	2.485
中日	広島	広島市民	12	10	2.485
中日	阪神	ナゴヤドーム	12	11	2.485

中日	阪神	甲子園	11	4	2.398
----	----	-----	----	---	-------

付表 2. 2009 年度ホームラン数

チーム	対戦チーム	球場	試合数	合計ホームラン数	オフセット
ヤクルト	横浜	横浜スタジアム	9	5	2.197
ヤクルト	横浜	神宮	13	13	2.565
ヤクルト	巨人	神宮	11	8	2.398
ヤクルト	巨人	東京ドーム	9	4	2.197
ヤクルト	広島	マツダスタジアム広島	12	11	2.485
ヤクルト	広島	神宮	10	11	2.303
ヤクルト	阪神	甲子園	8	5	2.079
ヤクルト	阪神	神宮	13	6	2.565
ヤクルト	中日	ナゴヤドーム	10	6	2.303
ヤクルト	中日	神宮	13	17	2.565
横浜	ヤクルト	横浜スタジアム	9	9	2.197
横浜	ヤクルト	神宮	13	10	2.565
横浜	巨人	横浜スタジアム	12	10	2.485
横浜	巨人	東京ドーム	10	16	2.303
横浜	広島	マツダスタジアム広島	12	9	2.485
横浜	広島	横浜スタジアム	12	10	2.485
横浜	阪神	横浜スタジアム	12	11	2.485
横浜	阪神	甲子園	11	6	2.398
横浜	中日	ナゴヤドーム	12	6	2.485
横浜	中日	横浜スタジアム	10	5	2.303
巨人	ヤクルト	神宮	11	14	2.398
巨人	ヤクルト	東京ドーム	9	12	2.197
巨人	横浜	横浜スタジアム	12	22	2.485
巨人	横浜	東京ドーム	10	12	2.303
巨人	広島	マツダスタジアム広島	12	12	2.485
巨人	広島	東京ドーム	9	8	2.197
巨人	阪神	甲子園	12	7	2.485
巨人	阪神	東京ドーム	11	18	2.398
巨人	中日	ナゴヤドーム	12	12	2.485
巨人	中日	東京ドーム	12	20	2.485
広島	ヤクルト	マツダスタジアム広島	12	11	2.485
広島	ヤクルト	神宮	10	13	2.303
広島	横浜	マツダスタジアム広島	12	6	2.485
広島	横浜	横浜スタジアム	12	17	2.485
広島	巨人	マツダスタジアム広島	12	5	2.485
広島	巨人	東京ドーム	9	8	2.197
広島	阪神	マツダスタジアム広島	10	3	2.303
広島	阪神	甲子園	9	4	2.197
広島	中日	ナゴヤドーム	12	4	2.485
広島	中日	マツダスタジアム広島	12	8	2.485
阪神	ヤクルト	甲子園	8	7	2.079
阪神	ヤクルト	神宮	13	8	2.565

阪神	横浜	横浜スタジアム	12	7	2.485
阪神	横浜	甲子園	11	9	2.398
阪神	巨人	甲子園	12	8	2.485
阪神	巨人	東京ドーム	11	15	2.398
阪神	広島	マツダスタジアム広島	10	3	2.303
阪神	広島	甲子園	9	15	2.197
阪神	中日	ナゴヤドーム	12	9	2.485
阪神	中日	甲子園	9	3	2.197
中日	ヤクルト	ナゴヤドーム	10	7	2.303
中日	ヤクルト	神宮	13	15	2.565
中日	横浜	ナゴヤドーム	12	10	2.485
中日	横浜	横浜スタジアム	10	18	2.303
中日	巨人	ナゴヤドーム	12	10	2.485
中日	巨人	東京ドーム	12	15	2.485
中日	広島	ナゴヤドーム	12	7	2.485
中日	広島	マツダスタジアム広島	12	10	2.485
中日	阪神	ナゴヤドーム	12	7	2.485
中日	阪神	甲子園	9	10	2.197

付表3. 2010年度ホームラン数

チーム	対戦チーム	球場	試合数	合計ホームラン数	オフセット
ヤクルト	横浜	横浜スタジアム	10	13	2.303
ヤクルト	横浜	神宮	10	12	2.303
ヤクルト	巨人	神宮	10	9	2.303
ヤクルト	巨人	東京ドーム	10	12	2.303
ヤクルト	広島	マツダスタジアム広島	12	8	2.485
ヤクルト	広島	神宮	12	13	2.485
ヤクルト	阪神	甲子園	8	8	2.079
ヤクルト	阪神	神宮	12	11	2.485
ヤクルト	中日	ナゴヤドーム	12	5	2.485
ヤクルト	中日	神宮	10	10	2.303
横浜	ヤクルト	横浜スタジアム	10	8	2.303
横浜	ヤクルト	神宮	10	9	2.303
横浜	巨人	横浜スタジアム	10	10	2.303
横浜	巨人	東京ドーム	10	9	2.303
横浜	広島	マツダスタジアム広島	12	8	2.485
横浜	広島	横浜スタジアム	12	10	2.485
横浜	阪神	横浜スタジアム	11	14	2.398
横浜	阪神	甲子園	9	11	2.197
横浜	中日	ナゴヤドーム	10	8	2.303
横浜	中日	横浜スタジアム	11	11	2.398
巨人	ヤクルト	神宮	10	14	2.303
巨人	ヤクルト	東京ドーム	10	14	2.303
巨人	横浜	横浜スタジアム	10	15	2.303
巨人	横浜	東京ドーム	10	19	2.303
巨人	広島	マツダスタジアム広島	12	29	2.485

巨人	広島	東京ドーム	9	15	2.197
巨人	阪神	甲子園	12	8	2.485
巨人	阪神	東京ドーム	12	25	2.485
巨人	中日	ナゴヤドーム	12	11	2.485
巨人	中日	東京ドーム	11	22	2.398
広島	ヤクルト	マツダスタジアム広島	12	10	2.485
広島	ヤクルト	神宮	12	11	2.485
広島	横浜	マツダスタジアム広島	12	7	2.485
広島	横浜	横浜スタジアム	12	12	2.485
広島	巨人	マツダスタジアム広島	12	11	2.485
広島	巨人	東京ドーム	9	5	2.197
広島	阪神	マツダスタジアム広島	10	6	2.303
広島	阪神	甲子園	9	7	2.197
広島	中日	ナゴヤドーム	12	4	2.485
広島	中日	マツダスタジアム広島	12	9	2.485
阪神	ヤクルト	甲子園	8	6	2.079
阪神	ヤクルト	神宮	12	20	2.485
阪神	横浜	横浜スタジアム	11	16	2.398
阪神	横浜	甲子園	9	10	2.197
阪神	巨人	甲子園	12	16	2.485
阪神	巨人	東京ドーム	12	15	2.485
阪神	広島	マツダスタジアム広島	10	19	2.303
阪神	広島	甲子園	9	8	2.197
阪神	中日	ナゴヤドーム	12	5	2.485
阪神	中日	甲子園	10	9	2.303
中日	ヤクルト	ナゴヤドーム	12	7	2.485
中日	ヤクルト	神宮	10	11	2.303
中日	横浜	ナゴヤドーム	10	8	2.303
中日	横浜	横浜スタジアム	11	10	2.398
中日	巨人	ナゴヤドーム	12	9	2.485
中日	巨人	東京ドーム	11	10	2.398
中日	広島	ナゴヤドーム	12	7	2.485
中日	広島	マツダスタジアム広島	12	15	2.485
中日	阪神	ナゴヤドーム	12	6	2.485
中日	阪神	甲子園	10	13	2.303

参考文献

- [1] Dobson, A. J. (2002) *An Introduction to Generalized Linear Models, Second Edition*, Chapman and Hall (田中豊他訳 『一般化線形モデル入門』 共立出版)
- [2] McCullagh, P. and Nelder J. A. (1989) *Generalized linear models, Second Edition*, Chapman and Hall, London.

- [3] Nelder, J. A. and Wedderburn R. W. M. (1972) Generalized linear models. *Journal of the Royal Statistical Society, Series A*, 135, 370-384.
- [4] SAS institute (2012) JMP ver.10.0.2 manual 『モデルと多変量』
- [5] 2009 ベースボール・レコード・ブック, ベースボール・マガジン社
- [6] 2010 ベースボール・レコード・ブック, ベースボール・マガジン社
- [7] 2011 ベースボール・レコード・ブック, ベースボール・マガジン社