

## 表計算ソフトを使ってみよう —プロ野球記録への適用例—

永井 恭一

医療学部 医療工学科 臨床工学コース  
nagai@toua-u.ac.jp

### はじめに

表計算ソフトはいろいろと便利な使いみちがあります。合計や平均などの簡単な計算を始めとして、その他いろいろとむずかしい数式が計算できます。数値以外に文字も入力できます。また、計算以外にも条件に合うものを選出したり、その個数を計測することもできます。データ数が多いときには目で見て探すのは大変ですし非常に便利です。その他、表を使って簡単にグラフをかくこともできます。

例えば、変数  $x$  に対して次式は正規分布を表すかなりむずかしい式ですが、このような関数も表計算ソフトで比較的簡単に計算できます。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} \quad (1)$$

ここで、 $m$  は平均、 $\sigma$  は標準偏差と呼ばれるものです。正規分布は理論的分布であり、グラフで

かくと釣り鐘状の形で表されます。自然界に存在するデータのばらつきぐあいを調べると、その多くが正規分布で近似できます。

図1は公称値 1.6k  $\Omega$  のカーボン抵抗 240 個について、その抵抗値を測定してグラフにしたものです。ヒストグラムと呼ばれています。1.6k  $\Omega$  と表示されていても、誤差を含んで作られていますので、測定するとこのように値に幅が出てきます。縦軸はそれぞれの抵抗値の個数ではなく、合計の数で割り算した割合 (%) で示しました。グラフをかくには表を作っておく必要がありますが、ここでは表は省略します。抵抗値は 1575 ~ 1592  $\Omega$  であり、中央近くの値が多く、中央から離れるにつれて左側、右側ともほぼ同じように少なくなっていることがわかります。

この図には棒グラフのほかに、(1) 式で求められる正規分布の曲線が重ねて示されています。このように違った種類のグラフを重ねて示すことも

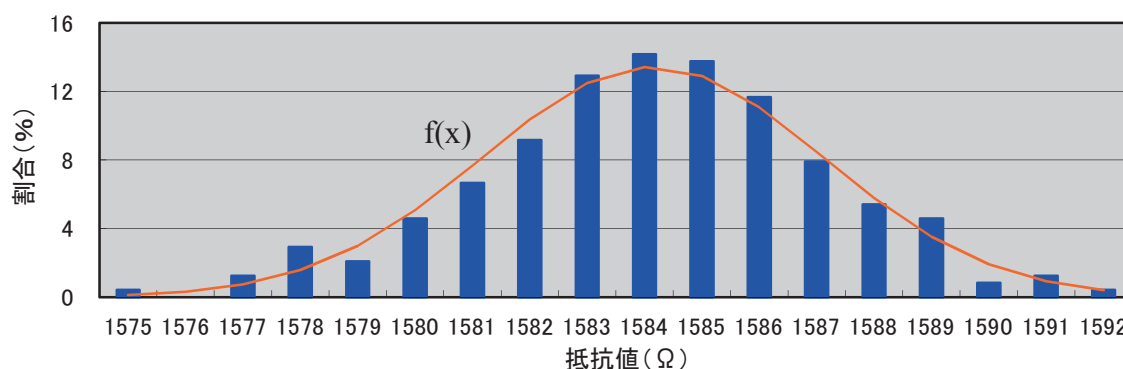


図1 カーボン抵抗の抵抗値の分布

できます。 $f(x)$ はある区間  $a \leq x \leq b$ での確率(図1では割合)を表すもので、具体的にはこの区間での曲線下の面積が確率になります。2つのグラフを比較するにあたり、(1)式の曲線下の面積を近似として  $x = 1575, 1576, \dots$ に対する  $f(x)$ の値  $\times 1$ の長方形の集合と考えることにします。このように考えると、2つのグラフを縦軸の値で比較することができます。両者はかなりよく対応していて、カーボン抵抗の分布が正規分布に近いことがわかります。

## スポーツについて

2020年の東京オリンピック・パラリンピックが決まり、日本が元気になったように思います。若い人の中には大きな目標ができた人もいるでしょう。また、大きなイベントとは関係なく普通にスポーツを行っている人も多くいるだろうと思います。スポーツは健康のためであったり、体を動かしてすがすがしい気持ちを楽しむというだけでなく、これがストレスの発散になったり、目標に向かって努力する気持ちを育んだり、仲間との協調性を身につけるなど精神的な面があることも知られています。あるアンケートによれば、子供にスポーツをさせたいかどうかの質問に85%以上が「はい」と答えています<sup>(1)</sup>。

一方、自分で行うスポーツのほかに、見て楽しむスポーツもあります。好きなプロスポーツを尋ねるアンケートをみると、だいたい1位はプロ野球です<sup>(2)</sup>。日本では野球が人気のスポーツといえます。若い人の中にはサッカーと答える人も多いようですが、年配の特に男性に根強い人気がある

ようで、広範囲な年齢でみるとやはり野球が1位となるのでしょう。

## プロ野球の楽しみ方

表計算ソフトはいろいろな方面に利用できますが、ここではプロ野球のデータに使ってみることにしました。プロ野球の楽しみ方には次のような3段階があるだろうと思います。キャンプ期間中には選手の動きを見てオーダーを予想したり、チームの順位を予想したりする楽しみがあります。これはどちらかという、「通」といわれる本当に好きな人の楽しみ方のようです。ペナントレース中は実際の試合を見る楽しみがありますし、ひいきのチームがあれば応援して一喜一憂することになります。これが一般的な見る野球の楽しみ方でしょう。ペナントレースが終われば、いろいろな記録を調べたり、まとめる楽しみがあります。これはややマニアックな人の楽しみ方でしょうか。ここでは表計算ソフトを使いますので、3番目の話になります。

## 平均打率

打撃について考えると、打率(=安打数÷打数)は最も基本的なデータと位置づけられます。表計算ソフトでA列に打者名、B列に打数、C列に安打数、D列に打率などとして表を作っておくと、例えばA2セルにある打者の打率はD2セルに「=C2/B2」と割り算の式を入力すると値が表示されます。また、次の行からはコピーによって値が簡単に得られます。

図2はセ・パ両リーグの、すなわち2リーグ制

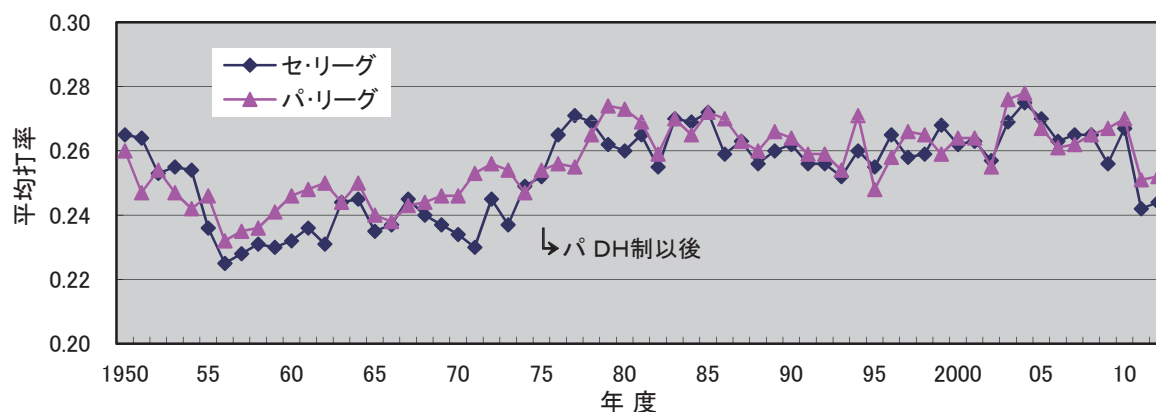


図2 平均打率の年度ごとの変化

以後のリーグ平均打率の各年ごとの変化を示したものです。一般にいくつかの項目があり、それらの数の大小を比べる場合には棒グラフが適していますが、1つまたはいくつかの数量の変化を表すときには折れ線グラフが適しています。セ・リーグ、パ・リーグで年度と平均打率の表を作っておくと、表の数値から簡単に折れ線グラフをかくことができます。グラフにすると、一目で打率の変化のようすがわかります。年度が変わり表のデータが増えた場合も、グラフをかき直す必要はなく、データの範囲を変更すればグラフが自動的にかきかわります。

図を見ると上下にかなり振れてばらつきがありますが、大まかな傾向として平均打率は1950年代始めから減少し、50年代中頃から70年代中頃まで低い値であり、その後上昇していますが、上がり続けることはなくばらつきの範囲内ではぼ一定です。また、2011年から統一球の導入によって低反発材の飛ばないボールが使用されて打率は明らかに低下しています。平均打率の最高はセ・リーグで0.275、パ・リーグで0.278です。どちらも2004年の記録です。

### 高打率の打者

プロ野球の歴史を振り返ると、初期には用具や球場の設備が整ってなかった時期がありました。また、いくつかの時期に飛ぶボールが使用されたこともありました。このほか環境の違いなどがあり、長い間にはヒットが出にくい時期やヒットが出やすい時期がありました。リーグ平均打率はこれを反映している数値といえます。

ここではリーグ平均打率に対する打者の打率の比を考え、これを「打率比」と呼び、以後使用することにします。この比は、打者がリーグの平均値に対してどのくらい高い打率を示したかを表していて、年度による好条件・悪条件に影響されにくい比率と考えられます。

打者の目標の一つとして、おそらく打率3割以上が挙げられると思います。さらに上をめざす場合には、首位打者を取ろうという気持ちもあるでしょう。ただ、首位打者は1年にそれぞれのリーグで1人ずつなので、誰でもというわけにはいかず非常に高い目標になります。ここでは各年のセ・パ両リーグの首位打者の打率を調べ、リーグ平均打率に対する比、すなわち首位打者の打率比を求めました。2リーグ制以後の首位打者の打率比は1.21～1.56の範囲の値でした。

図3はセ・パ両リーグを含めて、首位打者の打率比で分類した分布を示しています。すべての選手について打率比を求めると、1以下の値も出てきますし、前の図1に似た分布を示すと予想できます。しかし、図3では首位打者だけを対象にしましたので、図の打率比の中央で人数が最大になるのではなく、打率比の小さい側で最大になっています（人数も打率比の小さい側に偏っています）。図3で首位打者の打率比1.4以上は24%であり、この値は一流打者でも特に調子のよかったシーズンの記録といえます。首位打者の打率比1.4以上を2回以上記録しているのは長島、王、張本、イチローなどの7名で、超一流の打者の証とみることができます。

また、図3の首位打者はすべて規定打席に届い

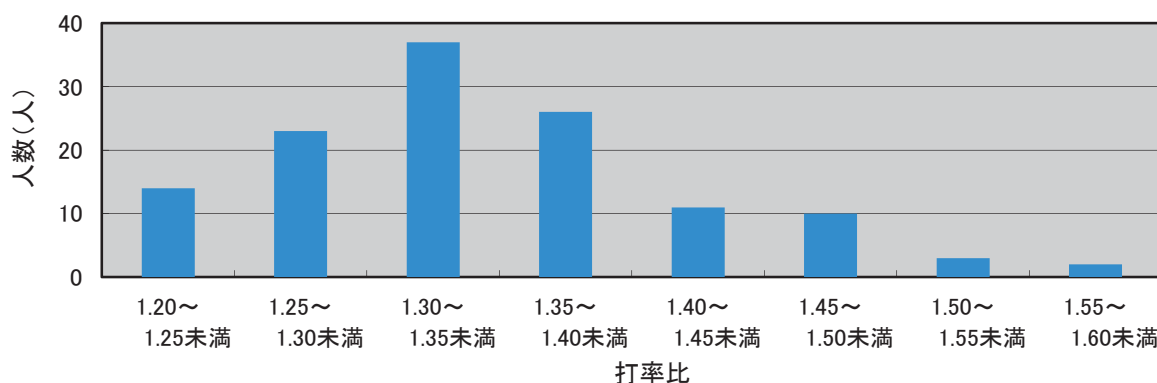


図3 首位打者の打率比の分布

表1 シーズン最高打率ランキング (2リーグ制以後)

| 順位 | 選手名   | 所属    | セ・パの別 | 年度   | 打率     | リーグ打率 | 打率比  | 修正打率  |
|----|-------|-------|-------|------|--------|-------|------|-------|
| 1  | バース   | 阪神    | セ     | 1986 | 0.389  | 0.259 | 1.50 | 0.378 |
| 2  | イチロー  | オリックス | パ     | 2000 | 0.387  | 0.264 | 1.47 | 0.370 |
| 3  | イチロー  | オリックス | パ     | 1994 | 0.385  | 0.271 | 1.42 | 0.358 |
| 4  | 張本 勲  | 東映    | パ     | 1970 | 0.3834 | 0.246 | 1.56 | 0.393 |
| 5  | 大下 弘  | 東急    | パ     | 1951 | 0.3832 | 0.247 | 1.55 | 0.391 |
| 6  | クロマティ | 巨人    | セ     | 1989 | 0.3781 | 0.260 | 1.45 | 0.366 |
| 7  | 内川 聖一 | 横浜    | セ     | 2008 | 0.378  | 0.265 | 1.43 | 0.359 |
| 8  | 川上 哲治 | 巨人    | セ     | 1951 | 0.377  | 0.264 | 1.43 | 0.360 |
| 9  | ブルーム  | 近鉄    | パ     | 1962 | 0.374  | 0.250 | 1.50 | 0.377 |
| 10 | 谷沢 健一 | 中日    | セ     | 1980 | 0.3694 | 0.260 | 1.42 | 0.358 |
| 11 | ローズ   | 横浜    | セ     | 1999 | 0.3685 | 0.268 | 1.38 | 0.347 |
| 12 | 落合 博満 | ロッテ   | パ     | 1985 | 0.367  | 0.272 | 1.35 | 0.340 |

ている打者の数です。というのも、首位打者については一つの特例があり、規定打席に届いていなくても首位打者になれる可能性があるからです<sup>(3)</sup>。すなわち、不足する打席数をすべて凡打したとして（規定打席に届くようにして）打率を計算したとき、これが規定打席以上の選手の最も高い打率より高ければ、首位打者として認められます。しかし、セ・パ両リーグでは一度もこのような例はありません。好調で高い打率を維持しているときに、けがなどで短い期間チームから離れる場合のように偶然の要素もありますので、なかなか該当するのはむずかしいでしょうが、野球ファンとしてはこのような首位打者の実現を見てみたいものです（二軍ではこの適用例があります）。

野球ファンの期待は日本で4割打者が出現しないのだろうかということでしょう。メジャーリーグでも1941年のテッド・ウィリアムズを最後に70年以上出てない大変な記録です。打率について歴代ランキング（2リーグ制以後）を記すと、表1のようになります<sup>(4)</sup>。3割8分台の打率を残すことはできていますが、その先の4割は高い壁のようです。この表を見てわかることは左バッターが多いということです。この表で右バッターは内川、ローズ、落合の3人です。打率を稼ぐには左バッターが有利であることが想像できます。

ここで、順位6位のクロマティは開幕から好調で96試合目まで打率が4割に届いていました。その後しだいに打率は下がりましたが、4割を維持していたときに規定打席数に届いていましたので、その後の試合に出なければ4割打者は誕生していたことになります<sup>(5)</sup>。しかし、チームは優勝

争いをしていましたし、助っ人として来日していたので残りの試合に出ないわけにはいきませんでした。最後は3割7分台の打率になりましたが、「幻の4割打者」ということになります。

また、表1にある修正打率は打率比に0.252をかけた値です。この0.252は広尾によれば<sup>(6)</sup>、日本プロ野球の全期間を通じた平均打率とのことです。ただし、この全期間とは1936年～2012年までのようです。ここで用いた修正打率は時代によって異なる打撃データを公平に判断するための1つの指標と位置づけられます。これを使うとリーグ平均打率が低いときの選手個人の打率は相対的に高くなりますし、リーグ平均打率が高いときには逆に相対的に低くなります。これである程度客観的な打率が得られます。修正打率で歴代ランキングを表し直しますと、1位は0.393の張本、2位は0.391の大下となります。0.387のイチローは修正打率では0.370で10位に下がります。この修正打率でも4割に届く選手はいません。

### 散布図による2つの量の関係

投手の代表的データとして防御率があります。防御率は投手が9回投げたときに何点取られたかを表す数値（比率）で、計算式は「自責点（投手の責任による失点）÷投球回数×9」です。防御率はタイトルの1つにもなっていて、これが低いということは点を取られにくい投手ということになります。ただし、エラーによる失点は自責点にならないので、二死からエラーで走者が出てその後打たれて何点取られても自責点にならないなど防御率にはいくらか問題点がありますが、ここでは打たれにくかったことに対する1つの目安と考

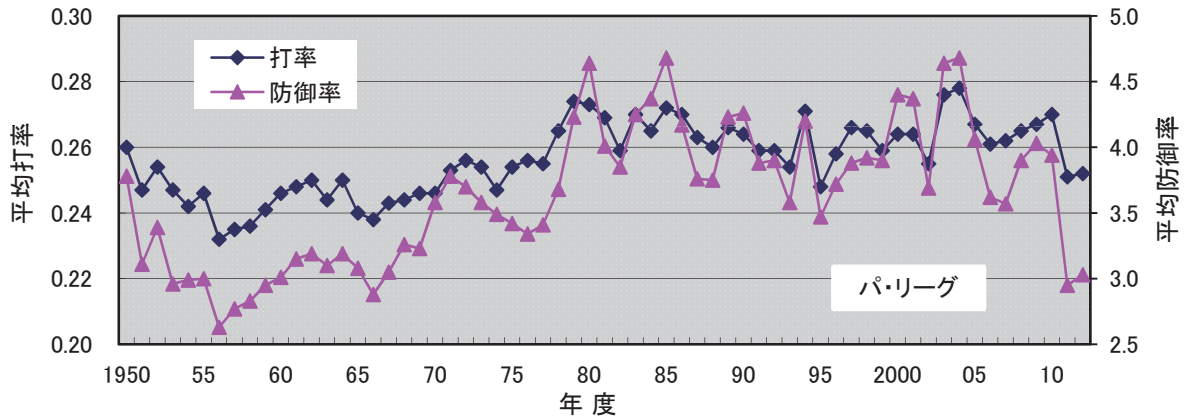


図4 パ・リーグの平均打率と平均防御率の年度ごとの変化

えます。

平均打率が高い年はヒットが多く出て、打撃戦が多かったものと思われます。また、平均防御率が高ければやはり打撃戦が多かったことになるでしょう。すなわち、打率が高いシーズンは防御率も高くなるだろうと予想されます。

1つのグラフで同じ横軸に対して、2つの異なる数量を取りたいとき、縦軸に第2軸を取ることができます。すなわち、第2軸に主軸と異なる目盛をつけることができます。図4はパ・リーグの平均打率と平均防御率の各年ごとの変化をグラフにしたものです。図2をかくときに年度と平均打率の表を作っていますので、この表に平均防御率を追加しておくこととなります。平均防御率の最高はセ・リーグで4.39（2004年）、パ・リーグで4.68（1985年、2004年）です。同じグラフで

打率と防御率の変化を表していますので、2つのデータの変化が比較できますが、図4ではこの2つの量が対応しているようであるという程度までしかわかりません。

図4は年度ごとの平均打率と平均防御率で、各年の変化が知りたい場合のグラフなので、横軸は等間隔で折れ線グラフでした。横軸にも縦軸と同様に数値的な意味を持たせて、打率と防御率を座標で点を取るように表すと、2つの数量の関係を知ることができます。パ・リーグについて横軸に平均打率、縦軸に平均防御率をとり、これでグラフをかいてみると図5のようになります。このグラフは散布図と呼ばれています。

図5を見ると、ややばらつきはありますが、直線関係が見られ、平均打率が高いと平均防御率も高い値となっていて両者はよく対応していることがわかります。このように2つの量の間にある種の関係があるとき、両者には相関関係があるといえます。また、このグラフでデータが狭い範囲で直線に近いほど強い相関であり、データ点にふくらみ（ばらつき）があるほど弱い相関ということになります。データ点が点在してどう見ても関係がなさそうときには2つの量に相関関係がないこととなります。

これは目で見て判断するのではなく、データの値を使って相関係数  $r$  という数値を求めることで判断できます。相関係数  $r$  は  $-1 \sim 1$  までの値を取り、絶対値で1に近いほど相関が強く、 $r$  が1から低下するにつれて、データはふくらみを持ち弱い相関から相関がない状態になります。図5

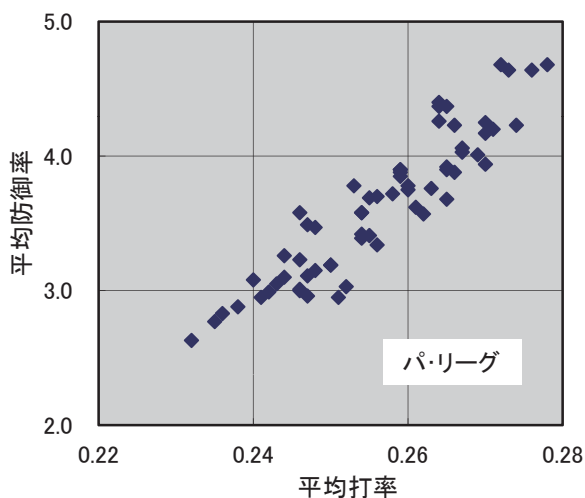


図5 パ・リーグの平均打率と平均防御率の関係

のパ・リーグの打率と防御率の関係では  $r = 0.92$  であり、図はありませんがセ・リーグの打率と防御率の関係では  $r = 0.95$  でした。

このように2つの量の関係を知りたいときには横軸、縦軸を数値で取り、データ点をその座標で表す散布図が適しているといえます。また図は省略しますが、平均打率とチーム1試合当たりの打点との関係を調べると、セ・リーグ ( $r = 0.94$ )、パ・リーグ ( $r = 0.92$ )とも良い相関関係がありました。

### 指名打者制 (DH 制) の影響

パ・リーグには1975年から指名打者制 (DH 制) が導入されました。この場合投手は打席に立たないので、平均打率がいくらか上がりそうな気がします。同様に考えると、DH 制により平均防御率も上がりそうです。

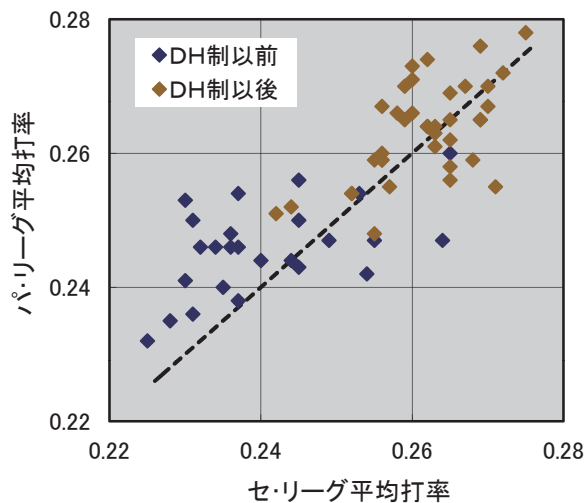


図6 セ・リーグの平均打率とパ・リーグの平均打率の関係

パ・リーグではDH 制以前と以後で平均打率に違いが出ているのでしょうか。前にもどり図2を見ても、DH 制以後の打率が以前の打率に比べて違いがあるかどうかははっきりわかりません。また、設備や環境面を考えると、DH 制のない74年以前と75年以後では74年以前の方が不十分であるように思われます。このほかにも様々な要因があり、75年を境とする打率を同じ条件で比べるのはむずかしく、平均打率についてDH 制だけの影響を調べるのはなかなか大変です。

この解決法として、セ・リーグの平均打率を使うことにしました。すなわち、セ・パの2つの打率を比べ、パ・リーグの平均打率がDH 制の影響を受けているかどうかを調べました。散布図が利用できそうであり、これを使った結果が図6です。中央の斜めの破線はセ・パの平均打率が等しいラインで、データがこれより左斜め上の領域にあればパ・リーグの打率が高く、反対に右斜め下の領域にあればセ・リーグの打率が高いことを示しています。

図6では全体的に、パ・リーグの平均打率がわずかに高い傾向にあるようです。DH 制以前と以後のデータを色違いで区別して表しましたが、この前後ではっきりとした違いは認められないようです。また、両者の関係を一次関数として近似式で表すこともできますが、相関係数  $r$  が低く、DH 制の前後で直線の傾きが異なるなどにより、ここでは数式を求めて検討することは行いませんでした。このほか、パ・リーグの平均防御率についても調べましたが、同様にDH 制の明確な影響は認められませんでした。

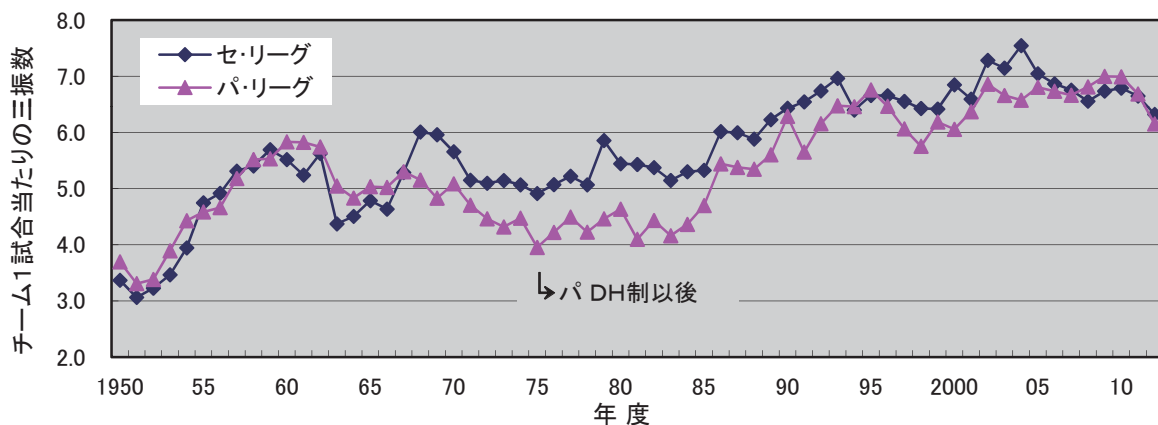


図7 チーム1試合当たりの三振数の年度ごとの変化

次に、三振について考えることにします。図7はセ・パ両リーグの年度ごとのチーム1試合当たりの三振数を示しています。概略的には三振数は1950年代始めから60年ぐらいまで大きく増加し、60年代中頃から85年ぐらいまではほぼ一定の範囲ですが、85年を過ぎるとまた増加していることがわかります。86年にストライクゾーンの改正が行われ、ストライクゾーンが低めにボール約1個分広がりましたので<sup>(7)</sup>、これ以後三振数は増えたのでしょう。また、90年以後はフォークボールを始めとして多種多様な変化球が使われる時代になり、三振が増えたと考えられます。

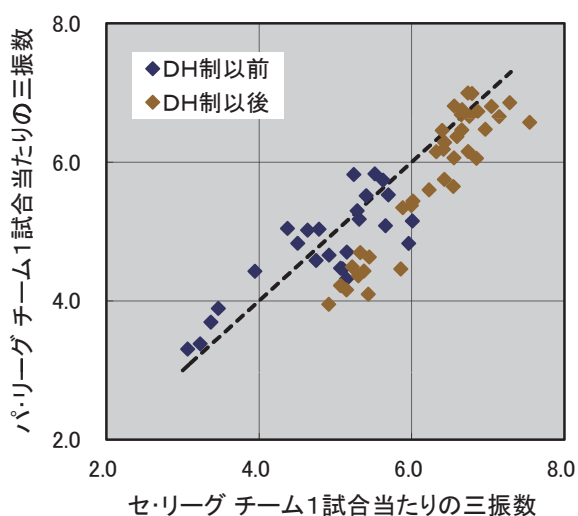


図8 セ・リーグの三振数とパ・リーグの三振数の関係

セ・パを比べると、60年代の終わりから90年代始めまでとその後95年から約10年間の時期に、パ・リーグの三振数が少ないことがわかります。DH制では投手が打席に立たないので、三振数は少なくなるだろうと予想されます。しかし、DH制以前にもパ・リーグの三振数が少ない時期があったり、DH制以後にパ・リーグの三振数が多い時期もありますので、図7でDH制によって三振数が減少したかどうかははっきり断定できません。

これも散布図で調べてみることにします。図8はチーム1試合当たりの三振数をセ・パで比べたものです。図6と同様、中央の斜めの破線はセ・パの三振数が等しいラインで1974年以前はセ・パの三振数はほぼ同じようです。DH制以後の

データは破線から右へ（セ・リーグが多くなる側へ）目視で0.5程度平行移動しています。すなわち、DH制以後はパ・リーグの三振数がはっきり少なくなっています。結論として、DH制の導入によって三振は1試合当たり0.5個くらい少なくなっていると考えられます。

DH制では1試合当たり3～4回、投手の打席が打者の打席に代わります。打者の打率は3割程度なので、1試合当たり3～4打席、投手が打者に代わってもヒットが1本増える程度で打率や防御率にはあまり影響しないと考えられます。一方、三振については1試合当たり3～4打席が投手から打者に代わると、1個弱程度は違いが出ることは十分考えられることで、図8で述べたDH制により1試合当たりの三振数が0.5個程度少なくなっているという結果は通常の試合の感覚と比べても妥当な値と考えられます。

### 最後に

ここでは野球の記録について表計算ソフトを使った一部の例を述べてみました。図2～図8の値には、日本野球機構『日本プロ野球記録大百科2004第4版』およびベースボール・マガジン社『2005ベースボール・レコード・ブック』～『2013ベースボール・レコード・ブック』のデータを使用しました。

表計算ソフトはこのほかにもいろいろと多くの使い方がありますので、使っていくうちに新しい使い方を知ることできるでしょう。具体的な例で使っていくと、覚えやすいものと思います。仕事や身近なところで使ってみると、便利な使い方が見つかり、生活の中にさらに幅が出てくるのではないのでしょうか。

### 文献

- (1) 「スポーツ」に関するWEBアンケート 株式会社ドウ・ハウス  
<http://www.dohouse.co.jp/news/research/20130328/>
- (2) 第22回「人気スポーツ」調査 一般社団法人中央調査会など  
<http://www.crs.or.jp/data/pdf/sports14>
- (3) 宇佐見徹也 (1993) 『プロ野球記録大鑑』 講

談社 p.263

- (4) 広尾晃 (2013) 『プロ野球なんでもランキング』 イースト・プレス p.24-30
- (5) 週刊ベースボール 7.3 (2000) ベースボール・マガジン社 p.15
- (6) 文献 (4) p.14-23
- (7) 週刊ベースボール 5.15 (2000) ベースボール・マガジン社 p.15-17