

## 評価をするということ



随 筆

石井博昭\*

Mathematical evaluation

Key Words : Ranking, Consensus Formation, MVP, Voting Data,  
Assignment of person

評価をすることは最近の“はやり”になっています。製品の性能評価や誤差評価などは工学では昔から行われていて、特に品質管理は日本のお家芸です。ウォルター・シューハート、エドワーズ・デミングなど欧米の学者が先駆者ですが、日本が世界をリードしています。

ところが、最近是人を評価するというのが“はやり”です。人間を評価することはものを評価するのと違って、大変難しいことです。人間が人間を評価できるのでしょうか。しかし人を人が評価することを考えましょう。まず評価する側の資質が問われねばなりません。たとえば、昨今、どこの大学でも学生による授業評価を行っています。私は大阪大学でも今の職場の関西学院大学でも授業評価は惨憺たるものです。往々にして成績のよい学生と真面目に授業を受けていない学生では評価は違うものです。評価をする権利は真面目に授業をうけている学生にのみ与えるべきとは思いますが、授業が下手でわからないから授業を受ける気がしないと言われればそれまでです。私が遠い昔に大学生だった頃はそんなに授業に出ていませんでしたが、最近出席をとる先生が多いせいか、授業に出ては来ますが、喋っていたり、携帯を弄っていたりで真面目に聞いていません。そのような学生が鬱憤晴らしにめちゃくちゃな評価をします。字が小さい、声が聞こえないとかで

すが、もともと後ろの方にいるのですから、前に出てくれば良いのです。

愚痴ではなく、本題に戻りましょう。数値的評価の方法にはいろいろあり、研究対象として次第に認知されていますが、論文を書くのが目的でもありませんし、その目的なら他にも優れた研究者、例えば、この随筆を書くよう勧めていただいた情報科学研究科森田浩教授などがおられますので、世で行われている代表的な評価について考えてみたいと思います。

議員や首長を選ぶのに各候補者を評価して1票を投じるのが投票ですが、このような政治的な場面以外にもよく使われます。大阪大学工学研究科には阪神タイガースファンも多いと思います。以下の表1は2003年のセントラルリーグのMVP決定の時の上位3選手の得票結果を示したものです。MVPは野球に関連する記者が1位、2位、3位と順位をつけて3名連記投票する。その結果を順位別に集計して、1位の各得票には5点、2位には3点、3位には1点の重みを与えてその重み和をスコアとして計算し、最高点の選手がMVPとして選ばれるという方法をとっています。1963年からこの形をとっています。表1では井川は $62 \times 5 + 75 \times 3 + 36 \times 1 = 571$ のスコアを得ています。矢野は $71 \times 5 + 57 \times 3 + 41 \times 1 = 567$ のスコアです。金本は $48 \times 5 + 23 \times 3 + 33 \times 1 = 342$ となります。従ってこの年は井川がMVPに選ばれました。すぐに気が付くと思いますが矢野選手は1位としての得票数が断然多いのですが、選ばれませんでした。もし、1位にもっと大きな重み、例えば6点を与えるとどうなるでしょうか。このとき井川は $6 \times 62 + 3 \times 75 + 1 \times 36 = 633$ で、矢野は $6 \times 71 + 3 \times 57 + 1 \times 41 = 638$ となり、矢野がMVPを得たはずですが、ここで、ポイントは5,3,1システムでは1位と2位の



\* Hiroaki ISHII

1948年7月生  
京都大学大学院工学研究科数理工学専攻  
博士課程満期退学(1976年)  
現在、関西学院大学理工学部 教授  
工学博士 オペレーションズ・リサーチ  
TEL: 079-565-7129  
FAX: 079-565-7129  
E-mail: ishiihiroaki@kwansei.ac.jp

表1 2003年度セントラルリーグ MVP 選出の時の投票結果

Candidate	First rank	Second rank	Third rank	Total score
Kei Ikawa	62	75	36	571
Akihiro Yano	71	57	41	567
Tomonori Kanemoto	48	23	33	342

表2 1999年度セントラルリーグ MVP 選出の時の投票結果

Candidate	First rank	Second rank	Third rank	Total score
Noguchi (Dragons)	59	72	41	552
Uehaa (Giants)	80	35	33	538
Sekikawa (Dragons)	55	62	43	504

重みの差と2位と3位の重みの差が同じになっていますが、6, 3, 1では1位と2位の重みの差が2位と3位の重みの差より大きくなっていることかと思えます。このようなケースは何度かみられます。表2は中日が1999年セントラルリーグで優勝した時のMVP選出の時のデータです。この時も1位に6点を与えると野口は611点、上原は618点になり同じことが起こります。

もちろん、評価に絶対的なものはありませんので、6点はおかしいという方もあるかと思えます。このように投票者が順位つき連記投票をして、各順位での得票の重み付き総和としてスコアを計算する例は多いと思いますが、MVPでは1人だけを選ぶのですが複数選ぶときには選び方にはいろいろな方法が考えられます。1つは1回で上位から同時に選出可能な人数を選ぶ方法です。もう1つは最高スコアから1人ずつ選ぶ、選ばれた人を除いて、再度連記投票をする。その結果から最高スコアの人を選ぶ、という方法で選出可能な人数回連記投票を繰り返すという方法です。

上記のように評価と人選は密接に関連しており、この場合は評価者を適切に選ぶこと自体から評価が始まります。大阪大学にとって大事な科学研究費補助金は多くの応募者の中から複数の研究者を選んで交付されるので、誰を審査員として選ぶというところから評価が始まり、いかに公平公正に評価するかというのがポイントかと思われれます。私自身はその

メカニズムはよくわかりませんが、評価がうまく機能しているように思います。評価のもう1つの役割は合意形成のためのものです。複数の評価者が複数の候補者を順位付けして、その結果を統合して全体として候補者に1つの順位付けを与える場合で、これは上記の科学研究補助金の決定にも使われているようです。これの特別で重要な応用例は研究室配属です。4回生になった4月、あるいは3回生の2、3月に学生が研究室を希望順位に従って連記投票したデータから、できるだけ希望に従って研究室に配属するというのが一般的かと思えます。研究室には上限があるのと、研究室の先生はできるだけ成績が良い学生が欲しいという制約があります。合意としての配属案は必ずしも学生の希望順と実際の配属先の単純な違いの和を最小にしているし、最小にすべき関数も明らかではありません。通常の手順では第1希望者がオーバーした研究室では先生がある基準、成績順あるいは関連した授業の成績順、熱意の大きさなどで上限までの学生を決めます。一方第1希望者が上限以下の研究室はそこまでの学生は決定で、残りの受け入れ上限はその分減らします。第1希望であふれた学生は、第2希望の研究室をチェックします。すなわち、第1希望で一杯になった研究室以外の研究室で、上限に空きがある研究室に第2希望で何人いるかをチェックします。上限をオーバーした場合は再び先生がある基準の下で上限まで受け入れます。上限以下の研究室は第2希望の学生を

すべて受け入れます。そして上限をその分減らします。次に残った学生の第3志望をチェックし、同じ操作を繰り返し、全ての学生が決まるまで繰り返します。ただ、この方法では、ある学生はほとんど行く気がない研究室に配属されることも往々にしてあります。成績がよくても、たまたま希望者が多い研究室を上位希望に選ぶと、このようなこととなります。学生だけでなく先生も取りたい学生の希望を書いて、選ぶべきだという意見もあるでしょう。これ1つとっても評価は難しいものです。

一方で通常投票は1回ですが、異なる角度から同じ候補者に順位付け連記投票をすべき場合があると考えます。例えばオリンピックのマラソン選手を選ぶ場合、実績、勝負強さ、スタミナ、実際に使われるコースに適しているか等複数の観点から順位付け連記投票してもらって、決めるのが良いのではないかと思います。同じタイプの選手ばかりではメダルを取るの難しいので、各観点での上位スコアから異なるタイプの複数選手を選ぶ方法がどうかと考えられます。高橋尚子選手が選ばれなかったのは記憶に新しいかと思えます。もちろん、どの観点から何人ずつを選ぶべきかは目的によって異なりますので、任意性ができます。最適なものはないように思えます。付け加えますと、実際のオリンピックの開催期日と代表として派遣する選手を決める期日とはかなりタイムラグがあるので、将来の伸びなどを見越したシナリオ解析を取り入れる必要もあります。男子フィギュアスケートの高橋が国内試合5位ながら、オリンピック選手に選ばれるなど少し不可解なことがスポーツの世界では多いので、評価をきちんとすることが必要かと思えます。体操やフィギュアスケートのように審査委員の採点で決まる競技が多いスポーツではどのような評価システムをとるかで結果が大きく異なります。この原稿を書いているときにソチ冬季オリンピックが行われていて、いつものことですが判定が疑問である競技がありました。人を順位付けすることの難しさを感じました。いつも評価方法は日本が不利にされてしまう場合が往々にして見受けられます。なにはともあれ、スポーツはこのように評価あるいは順位付けをどうするかの宝庫です。

また順位付け投票では投票者がどちらかに決定できにくい場合もあります。すなわちどちらが優れて

いるか迷う場合があります。例えば、2名とも甲乙つけがたく1位を決められないケースはどうかの場合です。この場合いわゆるtieが起こる場合です。幸い、得票数の重み和のケースでは、得票が分数でも使えるので投票は0,1ではなく、例えば2名A,Bについて同順位1位ではなく、候補者Aは1位として0.7票、2位として0.3票、候補者Bは1位として0.3票、2位として0.7票とする方法が考えられます。完全に同じ程度だとすれば、A、Bとの1位として0.5、2位として0.5とすればよいでしょう。tieが3人以上の場合などやtieが幾つもの順位である場合なども複雑にはなりますが分数票を許せば何とかなるかと思えます。

最後に順位ではなくカテゴリーで人を評価することを考えましょう。例えば野球で内野手（一塁手、二塁手、ショート、三塁手）の各ポジションにどの選手が適しているかです。よく適している、適している、適していない、全く適していない、の四つのカテゴリーでポジション毎に評価して、なるべく適しているポジションに割り当てたい状況を考えましょう。評価者は上記のように順番をつけにくい場合も多く、言語の方が評価する負担が少ないと思われます。ただ、これから結論を出すにはやはり数値化しなければなりません。このため、よく適している、適している、適していない、全く適していないをフェジ数としてメンバーシップ関数で表して、そのポジションに対する各選手の優勢度を計算します。この優勢度が選手をポジションに割り当てた時のメリットと考えて、メリットの総和が最大になるように選手をポジションに割り当てるという方法です。このように基本的な言語データから選んで評価をして、最終的に割り当て問題に帰着させるという方法は有効です。ただし、選手の方がポジションより多いのが現状なので、現実に対応するには工夫が必要です。このことは特に大学で人材を適材適所に配置するのに適していると思われませんが、いろいろな絡みがあり、困難な側面もあると思えます。

どちらにしても何にでも説明責任が問われる昨今、公正・公平さを保ちつつ、効率・効果的な人選の一助になればと思い取り留めのない話をさせていただきました。評価というものに関心をもっていたら望外の幸せであります。