

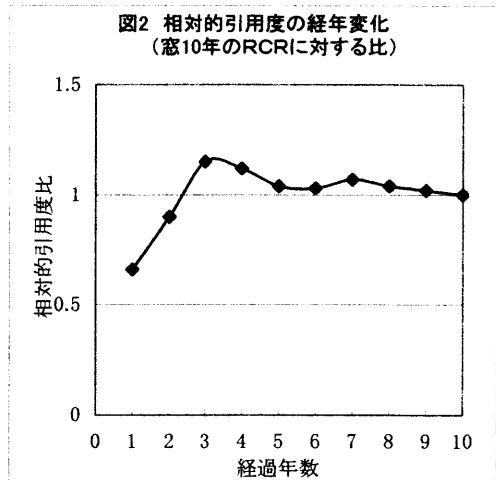
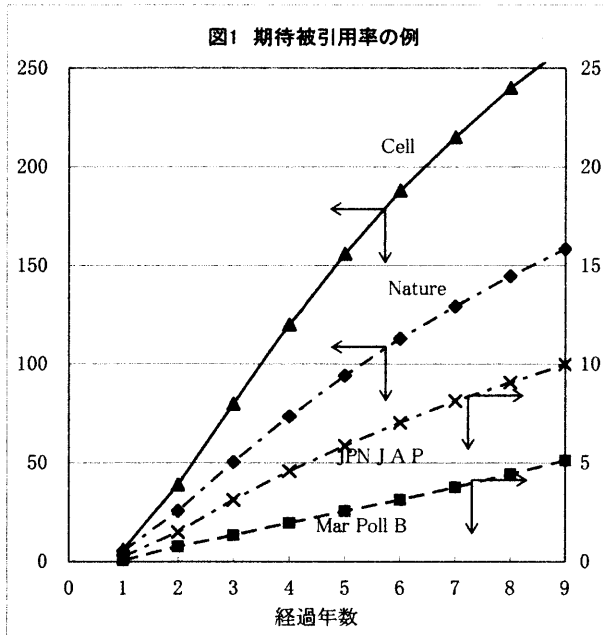
○大野博教（電中研）

1. はじめに

研究論文を主体とする研究業績評価における文献計量学的手法（以下BM法と略称）の利用は、ピア・レビューによる評価を補うもの、あるいはピア・レビューを支援するものとして、その存在価値が広く認められつつある。この場合、BM法では通常、学術誌への発表論文数をもって、研究の生産性を示す尺度として、また論文1篇あたりの被引用数、すなわち引用度をもってその論文のインパクトを示す尺度としている。しかしながら、論文の被引用数は専門分野によって、また学術誌の種類によって大きく異なり、この相違は場合によっては2桁にも達する。したがって、論文の被引用数の絶対値は、研究活動のインパクトを比較する際に、必ずしも適切な尺度とならない場合がある。この欠点を補うものとして、根岸らは、大学・研究機関の研究活動の評価の際、研究分野別の平均引用度に対する各大学の引用度の倍率を指数化したもの、すなわち引用度指数を算出した¹⁾。これは一種の相対的引用度であるが、研究分野別の平均引用度を求めるためにはかなりぼう大な作業量が必要となる。そこで、各論文毎に、その論文を掲載した学術誌の期待被引用率（Expected Citation Rate、以下ECRと略称）に対するその論文の被引用数の比、すなわち相対的引用度を求め、これを集計することによって論文群あるいは研究グループの研究活動のインパクトを求めることを試みた。なお、Science Citation Index（以下SCIと略称）の創始者、E. GarfieldやSCIの作成元のTomson ISI社のPendleburyは論文の被引用数を「ECRと比較する」とのべ、半定量的扱いにとどまっている^{2,3)}。したがって、ここに提案する「比」とは、広く認められた名称は無いが、便宜上、この比を相対的引用度（Relative Citation Rate、RCRと略称）と呼ぶこととする。

2. 相対的引用度の算出

相対的引用度（RCR）算出の基礎となる期待被引用率は図1に示すように学術誌によって大きく異なる。Mar Poll B誌では発行1年間のECRは0.057であり、1年以内に掲載された論文が1回引用されればRCRは17.5という突出した値となる。したがって、RCRを算出するためのデータ採取年数、すなわち窓（window）にはRCRが安定するまでの値をとる必要がある。この様子は図2に示す通りであり、一応5年でいどが適当と考えられている。なお、その根拠は示されていないが、Pendleburyも5年という数字を示している。



数字は愛媛大環境保護部の
主要論文 54 篇からの平均値

論文が複数の場合、特に種類が異なる学術誌の論文のRCRを総合して算出するためには、(1)複数の学術誌があたかも1種類の学術誌であるかのようにECRを算定する、(2)各学術誌毎にECRを求め、これらの算術平均をもって関連学術誌のECRとする、(3)各論文のRCRの算術平均をもって複数の論文の平均RCRとする、の3つの方法が考えられる。これらの方法をそれぞれ R_I 、 R_{II} 、 R_{III} として、以下に算出式を示す⁴⁾。

$$R_I = \frac{\sum_{j=1}^m \sum_{i=1}^n a_{i,j} / m}{\sum_{j=1}^m \sum_{i=1}^n A_{i,j} / \sum_{j=1}^m N_j} \dots \dots \dots (1)$$

ここに、 m : 論文数ならびにこれらの論文を掲載した学術誌の数 (重複を含む)
 $a_{i,j}$: j 番目の学術誌に掲載された論文が i 年目に受けた引用数
 $A_{i,j}$: j 番目の学術誌の発行年間の全論文の i 年目の総被引用数
 N_j : j 番目の学術誌の当該発行年間の総論文数

$$R_{II} = \frac{\sum_{j=1}^m \sum_{i=1}^n a_{i,j} / m}{(\sum_{j=1}^m \sum_{i=1}^n A_{i,j} / N_j) m} = \frac{\sum_{j=1}^m \sum_{i=1}^n a_{i,j}}{\sum_{j=1}^m \sum_{i=1}^n A_{i,j} / N_j} \dots \dots (2)$$

$$R_{III} = \sum_{j=1}^m (\sum_{i=1}^n a_{i,j} / \sum_{i=1}^n A_{i,j} / N_j) / m \dots \dots \dots (3)$$

上記の3種類の式の中、 R_I は最もオーソドックスな方法であるが、最も手間がかかる。 R_{II} はやや簡便であるが、複数の学術誌の間でECRに大きな差がある場合に、小さいECRが大きいECRにマスクされてしまう危うさがある。 R_{III} は最も簡便であるが R_I および R_{II} に対して実際に即したケース・スタディを試みる必要がある。3つの方法の試算の対象として、国のあるプロジェクトで得られた約30篇の論文を選び、発表年が同じ数篇ずつの論文群に分けて比較を行った。なお引用データ採取の「窓」としては、3年、4年および5年の3種類を選んだ。試算の結果から3つの計算式の間で相関を調べると、どの窓に対しても R_{II} と R_{III} の間で最も高い相関係数が得られた。そこで、算出の簡便さと論文の組み合わせに対する融通性から見て、 R_{III} が最も良いと考えられる。

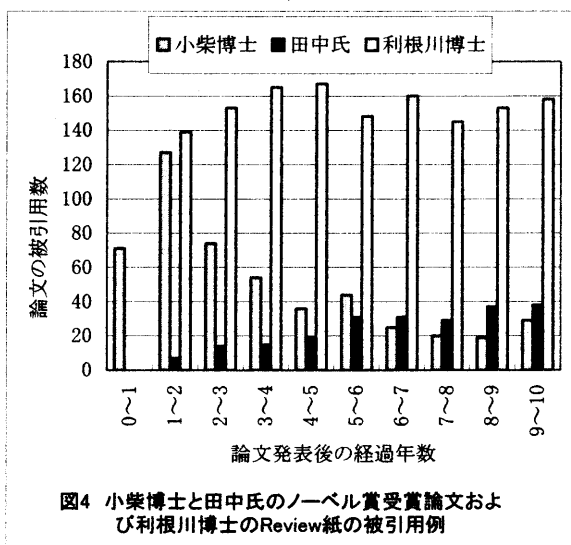
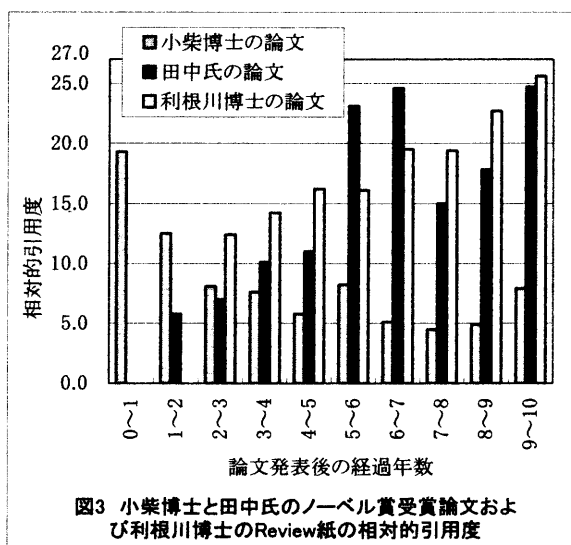
3. 相対的引用度の応用例

表1 相対的引用度の出現割合

RCR	割合、%
1以上	13.4
3以上	5.9
5以上	2.5
7以上	1.2
10以上	0.7

Nature誌1983、84、85年より収録

引用度では余り目立たない論文が、相対的引用度を求めると、驚く程きわ立ってくる場合がある。その実例として、昨年ノーベル化学賞を受賞された田中氏の状況を図3および図4に示す⁵⁾。相対的引用度が1ということは、その論文の被引用数が当該学術誌中の論文の被引用数の平均値に相当することである。しかしながら、被引用数の分布は寡占分布であり、分布に大きな偏りがあることに留意する必要がある⁶⁾。Nature誌を例にとると、表1に示すように、RCRが1以上の論文は、全体の中の10パーセント台に過ぎず、10以上の論文は1パーセント以下である。そこで、田中氏のようにRCRが20を越すということは、極めてまれな事象であると理解される。



4. おわりに

研究機関や研究グループの研究活動の評価にRCRを利用する場合に次のような利点がある。

(1) 専門分野に拘束されない

(2) 関係する研究者個人の各論文のRCRを集積すればよく、専門分野別の平均引用度を求める必要がない。すなわち、RCRの算出は自己完結型である。

なお、この場合、RCR算出のベースとなるECRは安定領域に入っていることが必要であり、このためにはデータ検索の窓を少なくとも5年にすることが望ましい。また研究者個人の業績評価にRCRを利用することも基本的には可能であるが、留意すべき点を挙げると次の通りである。

(1) 適切なクライテリアを設けること

(2) 利用目的を明確にすること

なお、これら2点については、今後ケース・スタディを積み重ねて、具体的に明示できるようにしたいと考えている。

稿を終えるにあたり、Nature誌に関する種々のデータはトムソンコーポレーション(株)の宮入暢子氏から提供されたことを付記する。

参考文献

- 1) 「大学ランキング, 2003年版」, 朝日新聞社
- 2) Garfield, E. : Current Contents, Sep.12(1994)
- 3) Pendlebury, D.A. : ISI 公開シンポジウム講演要旨, 10月8日(2002)
- 4) 大野博教 : 化学工学, 67, (8), pp.463(2003)
- 5) 大野博教 : 学士会会報, No.840, pp.115(2003-Ⅲ)
- 6) 小野寺夏生 : 学士会会報, No.838, pp.50(2003-I)