

学術論文の質について—被引用回数による分析

岩 本 直 樹 (香川大学工学部)

要 旨

新任採用、昇進、大学、研究所組織内部の教員、研究者、研究グループの生産性の評価等において学術論文リストは研究業績を評価するための基礎資料として不可欠である。中でも、学術論文リストの中の「論文数」は一目瞭然で、分かりやすい具体的数字であるため、しばしば「生産性」の指標とされがちである。本稿ではWeb of Science® 中の Science Citation Index Expanded™ を使い、論文の被引用回数を質の評価の基準に取り、サンプル集団について分析を行った。その結果、Web of Science® に採択されているようなある一定の基準を満たした「世界で最も権威と影響力のある高品質な学術雑誌」に掲載された論文と言えども、「論文数」は論文の質とは全く関係がない場合が多くあることがわかった。引用回数を使うに当たっての注意すべき点、研究者個人の評価と、その個人が属する研究グループ、組織の評価との関係、特に共著論文の取り扱い等における問題点等を議論する。

目 次

0. 要 旨
- I. はじめに
- II. 被引用回数検索
- III. 検索結果とその分析
- IV. 考察と引用データベース使用に当たっての注意点
- V. 結 語

I. はじめに

研究成果を発表しその内容を半永久的に記録する手段として学術論文誌が使われる。学術論文誌には、古くからある、紙に印刷された冊子形態のもの、およびWebで閲覧する電子ジャーナルと呼ばれる種類のものがある。それぞれの分野においてこのような学術論文誌に掲載された論文によって新しい知識が広められ、新たな発展に寄与する。このとき、査読付きの場合に限っても、学術論文誌に掲載される論文すべてが平等な価値を持っているわけでは決してない。それでは、多く出版される論文の中で「価値ある論文」、「優れた論文」といえるものはどのようなものであろうか？近年、大学や研究機関において日常茶飯事となった「研究評価」において、定量的に論文の価値を評価することは非常に重要である。¹論文の価値の評価はそれを生み出す研究者自身の評価にも直接つながる。例として、著者Aが一定期間（例えば1年間）に1編の論文を出版し、著者Bは同じ期間に10編の論文を出版したとする。このとき、1と10という分かりやすい数字が出てくるため「論文の数」を評価基準にすれば単純明快ではある。しかしながら、著者Bは著者Aに比べて研究者として10倍優れているということにはならないはずである。ではこのような場合どうやって比較するかという問題になる。もちろん、あらゆる場合に使える、公平かつ合理的な評価手段というのは存在しそうにない。しかしながら、ある程度論文の価値の定量的な指標となるのが論文の引用回数であると思われる。論文自体の電子化、それに伴う内容の分析、データベース化によってWebを使った引用データの検索は非常にやりやすくなってきている。本稿では引用データベースの中で最も充実している Web of Science[®] 中の Science Citation Index Expanded[™]² を使い、論文の被引用回数³ を質の評価の基準に取り、サンプル集団について分析を行い、その結果を議論する。

本稿の構成は以下のとおりである。まず、IIで使用したデータベースについて説明し、具体的検索方法について記述する。IIIでは検索結果の提示と分析を行い、IVで考察と引用データベース使用に当たった注意点、Vで結語を述べる。

II. 被引用回数検索

(1) 検索データベース

検索には Thomson Scientific 社の学術文献・引用検索データベース Web of Science[®] (<http://www.thomsonscientific.jp/products/wos/>) 中にある Science Citation Index Expanded[™] を使った。「Web of Science[®] は研究者が支持する世界で最も権威と影響力のある高品質な学術雑誌のみを厳選して収録したWebベースの学術文献データベース」が謳い文句である。

¹ 関連事項については以下を参照。岩本直樹 “アメリカの大学における教育・研究評価” 香川大学教養教育研究 第8号129-192 (2003).

² Science Citation Index Expanded, Institute for Scientific Information, Philadelphia, PA.

³ 「引用回数」といった場合「引用した回数」、「引用された回数」両方の意味がある。後者の場合「被引用回数」と書くべきであるが、以下の本稿では簡単のため、文脈でどちらか分かる限り単に「引用回数」を使う。

Web of Science は「全世界で9万冊以上発行されている学術雑誌、急速に増大する電子ジャーナルの中から最も影響力のある、本当に必要な情報のみを厳選して収録」していると記述されている。収録されている雑誌タイトル、数も毎年変化する。「毎年、3,000誌以上の電子ジャーナルを含む学術雑誌をレビュー、そのうち10—12%の雑誌を新たに収録」し「また、収録誌についても、収録当初のレベルが保たれているかどうかモニターして」いる。その中の Science Citation Index Expanded™ は「1900年以降、自然科学の150以上の分野から約5,900誌を収録」してある。データは毎週アップデートされる。⁴

(2) サンプル集団

工学部の一学科の構成員（2006年4月1日時点での教授全員）をサンプル集団とした。以下に具体的検索手順を示す。

(3) 著者名入力

著者名を入力してその名前で出版された論文を検索する。

(4) 本人確認

同姓同名、同姓同ファースト・イニシャルの研究者は所属を確認することで区別、選別した。あまりよく使われない名前の場合には本人のみが出てくるが、木村、田中、山田⁵ などよく使われる名前の場合には所属で本人を確認する必要がある。過去に所属をしばしば変えた著者に対しては履歴書（職歴）が必要になる場合がある。⁶

(5) 引用論文の検索

特定の著者の特定の論文についてそれを引用している論文を検索する。

(6) 自己引用の除外

引用している論文を検索する際、単著あるいは共著論文にかかわらず自分自身を含む著者の論文を引用している場合は数えない。

(7) 論文引用数の集計

(4)、(5)を特定の著者のすべての論文について行い、引用数の合計を求める。その際、単著、共著論文について別々に集計する。

⁴ Web of Science用のサーバは世界中に複数あり、どのサーバにアクセスするかによって最新データに関しては若干内容が異なる。これはアップデートの時期が異なることによる。

⁵ これらの名前は今回のサンプル集団には含まれていない。

⁶ これも今回は不必要であった。

(8) 時点、データ期間

引用データは2005年10月時点のもので、1980年以降の論文に関するものである。1979年以前の論文に関するデータは考慮されていないが、引用の傾向をつかむにはこれらで十分であると考えられる。特にここで扱ったサンプル集団では25年以前のデータの影響は極めて少ないことが分かる。Web上のデータは毎週更新されている。例えば、1年後2006年10月の引用データでは2005年10月以来新たに出版された論文を引用したものが加わり、さらに、2005年10月までに既に出版されていた論文の2006年10月までの1年分の引用データが加わる。したがって、分析の条件を明示するためにはデータ検索時点の日付が必要となる。さらに、データ保存のために、hard copyを残しておく必要がある。

Ⅲ. 検索結果とその分析

表1に検索結果を示す。以下それぞれの列についてグラフ化した。

表1： 検索結果。列は左から(1)著者番号、(2)論文数、(3)筆頭著者論分数、(4)筆頭著者論分数と全論分数の比、(5)論文引用回数、(6)論文1編あたりの引用回数、(7)筆頭著者論文引用回数、(8)筆頭著者論文1編あたりの引用回数を示す。

(1) 著者番号	(2) 論文数	(3) 筆頭著者 論分数	(4) 筆頭著者 論分数/ 全論分数	(5) 論文引用 回数	(6) 論文1編 あたりの 引用回数	(7) 筆頭著者 論文引用 回数	(8)筆頭著 者論文1編 あたりの 引用回数
1	41	8	0.195	893	21.8	109	13.6
2	32	8	0.25	484	15.1	137	17.1
3	35	15	0.429	77	2.2	13	0.867
4	27	3	0.111	103	3.81	11	3.67
5	90	46	0.511	695	7.72	263	5.72
6	26	20	0.769	652	25.1	636	31.8
7	23	8	0.348	80	3.47	38	4.75
8	16	4	0.25	25	1.56	11	2.75
9	39	7	0.179	567	14.5	260	37.1
10	46	27	0.587	607	13.2	426	15.8
11	35	6	0.171	514	14.7	186	31
平均	37.3	13.8	0.346	427	11.2	190	14.9

図1は各著者(1-11)について、SCIに掲載されている論文数の分布である。最低16、最高90、平均37.3で、最高と最高の比は5.6である。ここに、著者4の27編の論文中23編(85.2%)は著者7との共著である。一方、著者7の23編の論文のうち23編(100%)は著者4との共著である。

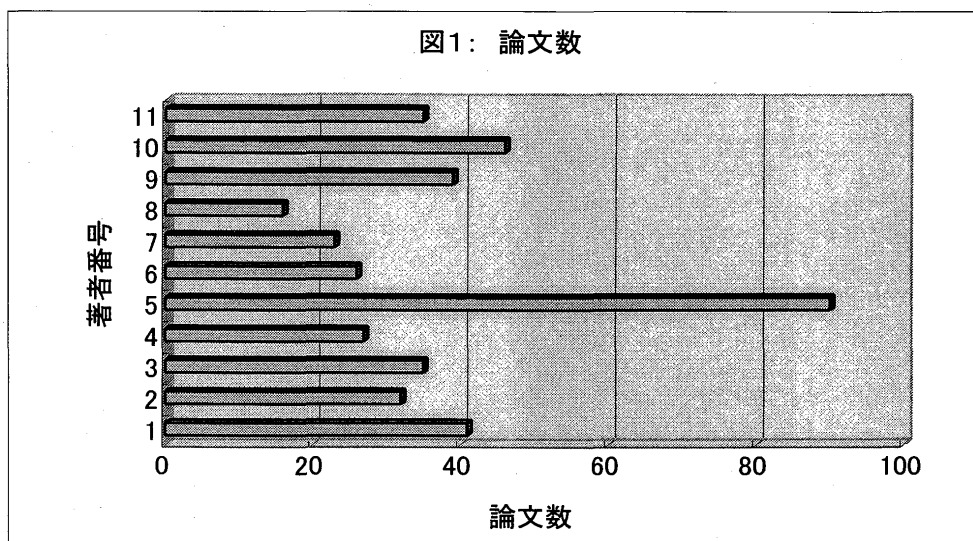


図2は各著者（1—11）について筆頭著者としての論文数の分布である。最低3、最高46、平均13.8、最高と最低の比は15.3と、一桁を超え、図1に示した全論文数の場合の比の2.7倍になる。ここで注目すべきは筆頭著者としての論文が3編しかない著者がいるということである。共同研究をしているもののほとんどの場合、筆頭著者とはならないという、特徴的な共同研究形態をとっているものと思われる。また、筆頭著者としての論文数と全論文数の比は各著者によって異なる。ここで扱ったサンプル集団では最低0.111、最高0.769、平均0.370、最高と最低の比6.93である。つまり、9編中1編の割合でしか筆頭著者として論文を出さない著者もいれば、4編中3編以上の割合で筆頭著者として論文を出す著者もいる。

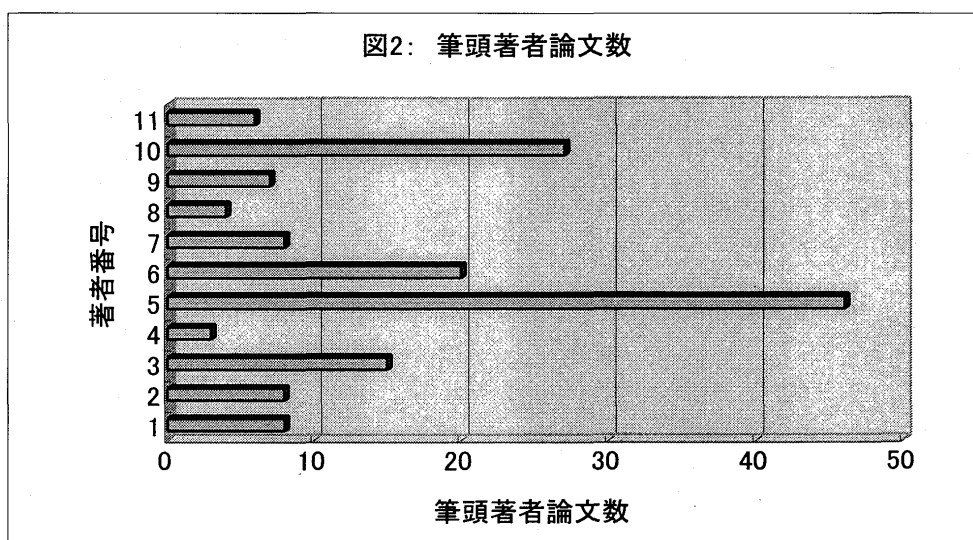


図3は図1に示した各著者の全論文に対する引用回数を示す。最低25、最高893、平均427で、最高と最低の比は35.7と大きく開きができる。これは、量と質とを掛けた（筆頭著者であるか否かを問わない場合の）全論文数に対する評価と見なせる。単に論文の数だけでは差が5.6倍あったが、ここでは

35倍を越える差が出てくることに注目したい。

とくに著者3, 4, 7, 8の研究は広く引用されているとは言いがたい。つまり、国際的に広く認められているとはいえないであろう。

著者を何人か比較すると注目すべき点が明らかになる。図1の論文数において著者2と著者3はそれぞれ32編、35編とよく似ている。しかしながら、図3の論文引用回数になると、著者2が484回に対して著者3は77回と6.3倍の差が出てくる。

同じようなことが著者4と著者6についていえる。論文数では著者4が27編、著者6が26編とよく似ている。ところが論文引用回数になると著者4の103回に対して著者6は652回と6.3倍の差が出る。論文リストのみを見る限りこのようなことは分からない。

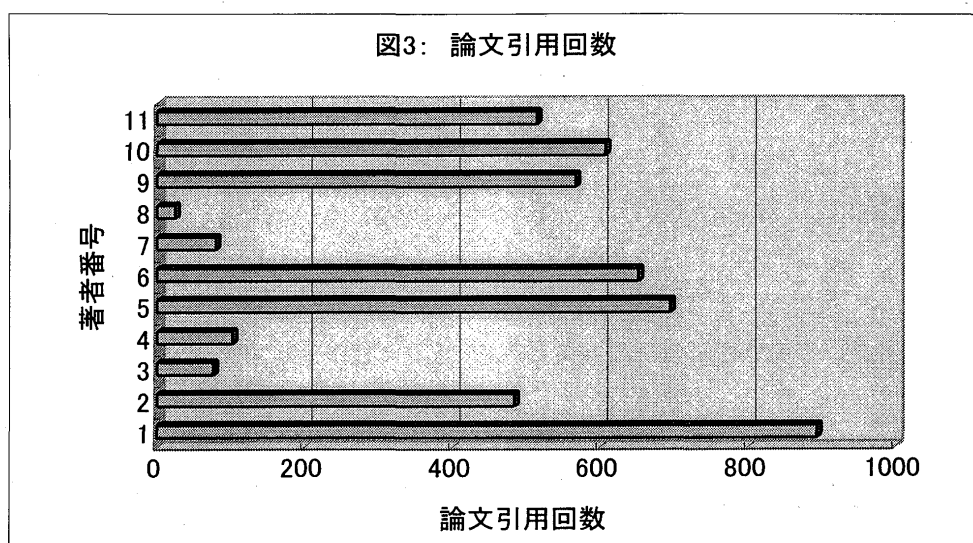


図4には各著者に対して筆頭著者としての論文の引用回数を示す。最低11、最高636、平均190、最高と最低の比57.8である。図3の(筆頭著者であるか否かを問わない場合の)全論文の引用回数に対して、筆頭著者としての論文の引用回数ではさらに開きが出るのが分かる。ここでも著者3, 4, 7, 8が目立って少ない。筆頭著者としての論文の引用回数が10回程度では、それぞれの分野での国際的な知名度には疑問があると言わざるを得ない。

先に比較した、論文数においてよく似ている著者4と著者6は、図4の筆頭著者としての論文の引用回数ではそれぞれ11回、636回と実に57.8倍という大きな差が出る。

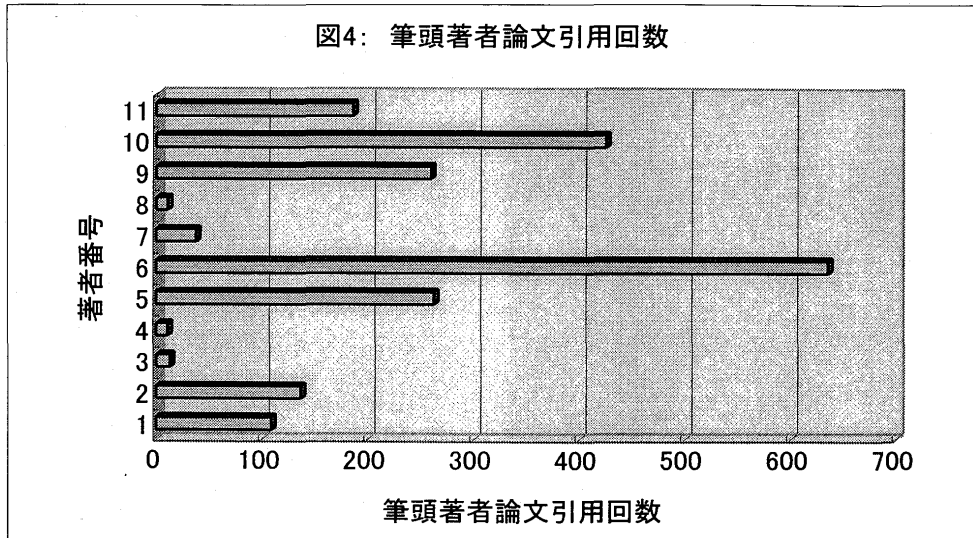


図5は各著者について（筆頭著者であるか否かを問わない場合の）論文1編あたりの平均引用回数を示す。最低1.56、最高25.1、平均11.2で最高と最低の比16.1である。これは1編あたりの（筆頭著者であるか否かを問わない場合の）論文の価値といえる。ここでも著者3,4,7,8、について図3、4で述べたことが同様に当てはまる。

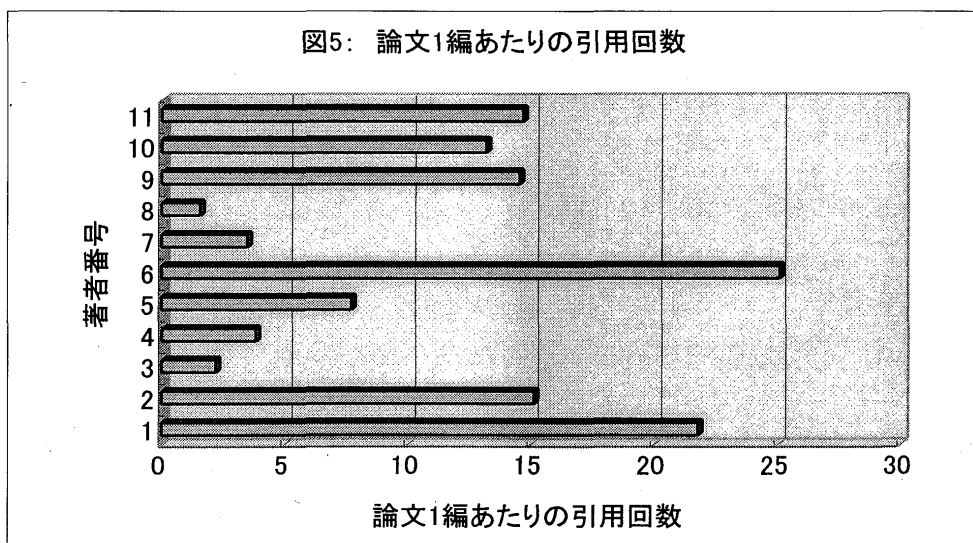
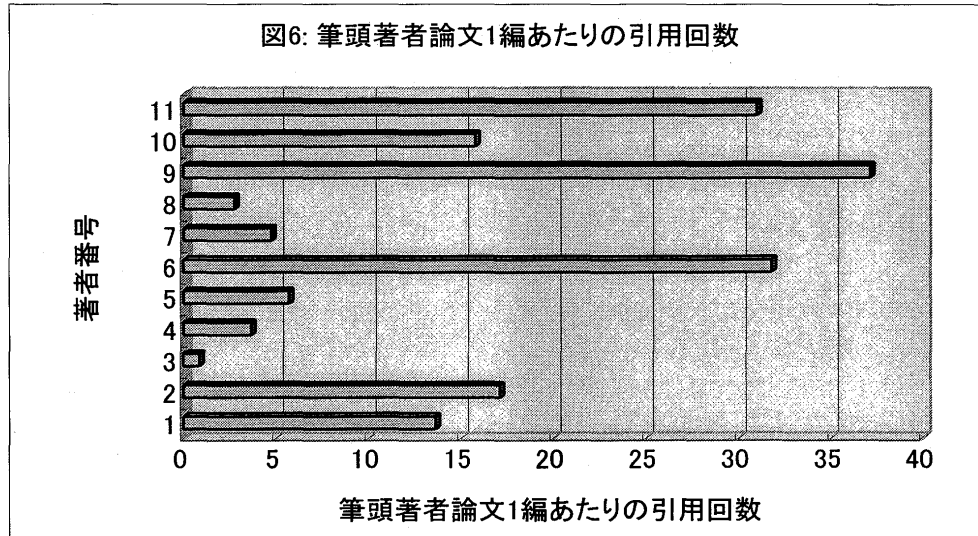


図6は各著者について筆頭著者としての論文1編当たりの引用回数を示す。最低0.867、最高37.1、平均14.9で最高と最低の比37.1である。これは筆頭著者としての論文1編あたりの価値といえる。



IV. 考察と引用データベース使用に当たっての注意点

まず、Ⅲで行った検索、分析において共著論文に関する以下の問題点を指摘しておく。

(1) 共著論文における各個人の寄与

共著論文で各著者をどう評価すべきかということはそれだけで難しい問題である。論文の寄与の程度によって評価すべきであるが、それがいつも定量的に分かるとは限らない。

(2) 研究グループ、組織の生産性の評価

研究グループ、学科、学部などの組織の構成員で共同研究がなされ共著論文が書かれ、これらの組織の生産性を考える場合、共著論文の取り扱いには特に注意を要する。

例えば、ここに10人で構成される研究グループあるいは組織があったとする。そして、ある年に10人の共著で10編の論文を書いたとする。この場合、各個人の論文リストには10編ずつの論文がその年には載せられる。しかしながら、この研究グループ全体で出した論文は10編×10人=100編ではなく、あくまで10編である。これは組織として平均一人あたり1編書いたことに相当する。

これに対して、このグループに属さない単独の研究者が単著でその年に1編の論文を書いたとする。さて、この年、このグループに属する研究者一人が共著で書いた論文の数10編と単独の研究者が単著で書いた1編の論文とはどう比べればよいのであろうか？前者が数では10倍の生産性を持つといえるのであろうか？単著で1年間に1編書く研究者が10人集まればその集団は1年間に10編の論文を出し、10人で構成される研究グループと生産性において同等である。もちろんこれらの論文の質は同じと仮定する。したがって組織としての生産性という観点からは、研究グループ構成員一人当たりの生産性の評価には共著論文の数（10編）を共著者数（10人）で割ったもの（1編）が評価材料として妥当ということになる。

このことは、研究グループの構成員の所属が複数の異なった組織（大学）である場合についても

同様に扱うべきであろう。例えば、A大学から3人、B大学から3人、C大学から4人の全部で10人から構成される研究グループの出す論文は、このルールを当てはめれば、全著者数（10人）で割った数を各人の出版論文数とするのが妥当となる。

Ⅲで分析した結果を使って具体例を示そう。すでに著者4の27編の論文中23編（85.2%）は著者7との共著であり、著者7の23編の論文のうち23編（100%）は著者4との共著であることに触れた。著者4と著者7をグループと考えたときこのグループの生産性はどうか？まず、論文数は27編+23編=50編とはならず、27編である。著者7のすべての論文は著者4との共著であるためグループとして見たとき独立な論文とはならない。次に引用数はどうか？著者4が103回、著者7が80回引用されている。ここでもグループとして見たとき引用回数は103回+80回=183回とはならず、同じ理由から103回である。著者個人個人で比べると引用回数は著者4、7とも著者3の77回に比べて多い。しかしながら、グループの一員として見たとき二人で103回の引用回数は一人当たり51.5回で著者3より少ないことになる。

このように、同じ組織内で行われる共同研究の生産性の評価には特別の注意が必要である。また、異なる組織間で行われる共同研究においても原理的に同じことが言える。ある組織の著者Aと別の組織の著者Bが共同研究を行う場合、論文数はそれぞれの組織で独立に数えられるがグループとしての生産性はそれぞれの論文数、引用回数の和とはならないからである。

次に引用回数を分析するに当たって一般的にいくつかの注意すべき点を挙げておく。

(1) 真価が理解されるまでに時間がかかる場合があること

出版されると直ちに引用され始める論文もあれば、相当時間がたった後はじめて引用が始まる論文もある。分野、テーマによって引用され始めるまでにかかる時間はまちまちである。今流行のテーマであれば引用という形での反応は比較的早いであろう。一方、誰もやらないことをやった場合、直ちに理解され引用されるとは限らない。また、一般的に、出版されて間もない論文は引用回数では評価しにくい。後にノーベル賞が授与されることになる研究論文であっても、出版当初誰も引用しなかったという例がある。それは、多くの可能性の中の一つについて書かれた論文であったためである。いったんそれが正しいことを証明する実験が出ると数多く引用されるようになった。このように、出版されて引用されるまでに時間がかかる場合がある。一方、分野、テーマによっては、出版前からpreprintの形で引用される場合もある。

(2) 分野の研究者数

研究者数が多い分野ではそれだけ引用される可能性が高い。全世界で研究者数が3人の分野ではすべての研究者に引用されても（全員が1編ずつ論文を出し、その中で引用するとして）3回となる。一方、研究者数100人の分野では、同じ条件下で100回引用される。

(3) 分野で出版される論文数

研究者一人当たりの論文数が多い分野ほど引用される可能性が高い。例えば、ある分野で研究者一人当たり年間平均論文1編を出版し、別の分野では研究者一人当たり年間平均30編出版するとす

る。この二つの分野で研究者数が同じ場合、当然後の方が引用回数は多くなるはずである。

(4) 間違っただ論文への批判

明らかに間違っただ内容の論文が出版され、一斉にその誤りを指摘する論文が出ることもある。この場合、学問的に価値のない間違いであれば、いくら引用される回数多くてもその論文に価値は見出せない。そのような種類の引用かどうかは、引用し、誤りを指摘するほうの論文を読めば分かる。ただし、簡単に決着がつかないとき、例えば、間違いだと主張する論文のほうが間違っただいる可能性がある場合、評価は単純にはできない。また、間違っただ論文でも、ある研究を刺激する結果になれば価値はあるかも分からない。⁷

(5) 最小出版可能単位

一連の測定結果を個々に短い複数の論文として出版するか、まとめて長い1編の論文として出版するか。1ページの論文を10編出すか10ページの論文を1編出すか。後者のほうが被引用回数は10倍となる。

(6) データベースと学術論文

実験手法が同じで試料、対象が数多くある場合、実験結果は学術論文誌に論文として出版するよりもデータベースに入れるほうが適当な場合がある。化学の分野で多数の化合物などを扱う場合にあてはまる。無数の星を研究対象とする天文学では「星の数ほど論文が書ける」とよくいわれる。

最後に、Science Citation Index ExpandedTM (SCI) 特有の問題点について述べる。

(1) 論文リスト→SCIのリスト

ある著者の論文リストに載っているすべての論文が必ずしもScience Citation Index ExpandedTM (SCI) の分析対象とはなっていない。Science Citation Index ExpandedTMは「1900年以降、自然科学の150以上の分野から約5,900誌を収録」している。ここに、第一段階のフィルターがある。たとえ頻繁に引用される論文であってもそれが掲載されている論文誌がSCIに収録されていないことには引用データは出てこない。したがって、ある著者の論文リスト中どれだけの論文がSCIに採択されているかということはそれだけで一つの評価となりうる。ただし、この評価基準はThomson Scientific社が設定したものであり、ユーザーが自由に設定することはできない。

(2) 日本語の学術論文誌と論文

SCIには日本語で書かれた論文は例外を除いて掲載されない。その例外とは、英語で書かれた論

⁷ ノーベル賞を2回受賞したJohn Bardeen先生があるとき、ある効果を理論的に予言した。多くの実験家はその効果を見つけるべく実験を行った結果、確かに新しい効果が実験的に見つかった。しかし、その理論は間違っただいて、たまたま別の効果を実験家たちが見つけたのだと後でわかった。偉い先生が論文を出すと、それが間違っただいても、分野の発展に寄与することがある。

文（もちろんその論文誌がSCIに収録されていることが条件である）が日本語で書かれた論文の雑誌名をアルファベットで引用した場合に限られる。したがって、理学系では通常ありえないが、工学系の分野によっては日本語の学術雑誌が主たる研究成果発表の場となることがありうる。このような場合、SCIを使った引用回数による論文の評価はできない。

(3) SCIに収録されている論文誌は毎年変わる

SCIに収録する論文誌のタイトル、数は毎年変化する。Web of Science全体では「毎年、3,000誌以上の電子ジャーナルを含む学術雑誌をレビュー、そのうち10—12%の雑誌を新たに収録」し「また、収録誌についても、収録当初のレベルが保たれているかどうかモニターして」いる。例えば、著者Aが2000年、2001年にそれぞれ論文誌Xに論文を出したとする。2000年の論文は論文誌XがSCIに収録されているため検索して出てくるが、2001年の論文は論文誌XがSCIに収録されていないので検索しても出てこないというようなことが起こりうる。

V. 結 語

(1) 1編の論文の価値

比較の結果、このサンプル集団では論文1編当たり平均で25回引用された著者もいれば、1.6回引用された著者もいて、相当引用回数のばらつきがある。つまり1編の論文に対する平均引用回数で16倍の個人差があることが分かった。

筆頭著者としての論文では、1編当たり平均37回から0.87回のばらつきが見られた。これは1編あたり40倍以上の個人差があるといえる。したがって、この場合、著者9の1編の論文の引用回数は、著者3の40篇の論文を合わせた引用回数にほぼ等しいといえる。従って、価値判断の基準を論文の引用回数にとると著者9と著者3の1篇の論文の価値には40倍の開きがある。著者9の論文リストと著者3の論文リストを単純に見比べるだけではこのことは全く分からない。

(2) 共著論文と研究グループ、組織の生産性

具体例を挙げて研究グループの生産性（論文数、引用回数）について論じた。グループの生産性は構成員の各々の生産性の和とはならないことが分かった。グループ構成員全員で共著論文を出す限り、2人のグループも10人のグループもグループ全体としての生産性は構成員一人の生産性に等しい。したがってグループ構成員一人当たりの生産性は各個人の生産性を構成員数で割ったものとなる。

(3) 研究評価の基礎データ

Ⅲで行った分析は各著者あるいは組織全体の研究評価のための基礎データとして十分使えると思われる。定期的に（例えば毎年）行えば、各著者あるいは研究組織全体の引用回数の時間変動も追える。

(4) 論文の数と質

本稿で明確になったように、研究評価が論文の数のみでなされるとしたらこれ以上に正当な評価から程遠いものはない。学生による授業評価も同様であるが、生のデータはあくまで生のデータである。それらを十分に吟味して初めて評価に使える判断材料となりうる。

(5) 総合評価

本稿で取り扱ったのは論文引用回数という観点から見た論文の質の評価とその著者の研究評価である。もちろん教員評価には競争的外部資金獲得状況などの研究評価、教育活動の面での評価、学内外の委員会などの管理運営活動などとともに総合的に評価されるべきである。このように考えたとき、幸いにも、引用回数については決して多くはなかった著者3, 4, 7, 8は評議員(1人)、学部長選挙立候補者(1人)、学科長(2人)、各種学部委員会委員長(複数)、学部運営委員(複数)、学科人事委員(複数)など全学、学部、学科管理運営において多岐にわたる重要な貢献をしてきており、総合評価においては全く問題はないと思われるということをつけ加えておきたい。

(6) 政策決定の判断材料として

いろいろ不完全な点はあるものの定量的に論文の質を評価する方法の例として引用回数を取り上げ、議論してきた。これは単なる「論文の数」あるいは組織構成員による「多数決」などに比べはるかに客観的で合理的な方法であると思われる。したがって、新規採用、昇進などの人事において候補者を評価するとき、研究者のグループ、組織の評価、そのほか、予算、人材などの資源配分の際に判断材料の一つとして使えるであろう。

(7) 個人情報

本稿では必要もなく、それが目的でもないので著者名に実名を出すことはしていない。ちなみに、仮に実名入りであったとしてもこれは法的に保護の対象となる個人情報ではない。なぜならこれらの情報はWebに公開されていて、著者名さえ知っていれば誰でも検索できるからである。言い換えると、論文を出版(公開)した時点で著者名は公開され、さらにその論文を引用した論文も公開され、となるからである。

謝 辞

The University of Toledo, Carlson Library, Electronic Information Services で電子情報検索に関して支援していただいた Professor Wade M. Lee に感謝します。

すると、もう、どんぐりどもが、くちぐちに言いました。

「いえいえ、だめです。なんといったって、頭のとがっているのがいちばんえらいのです。」

「いいえ、ちがいます。まるいのがえらいのです。」

「そうでないよ。大きなことだよ。」がやがやがやがや、もうなにがなんだかわからなくなりました。

宮沢賢治(1896-1933)　どんぐりと山猫(1921)