

第6回 SPARC Japan セミナー 2008 「IF を越えて—さらなる研究評価の在り方を考える」

研究評価・雑誌評価のためのビブリオメトリックス指標：現状と課題

根岸 正光
国立情報学研究所教授、SPARC Japan 運営委員長

要旨

インパクト・ファクター(IF)は学術雑誌の評価指標として有用であり、よく参照されるが、多くは購読雑誌の選定という本来の用法ではなく、研究者評価のための個別論文の評価指標として利用される。本講では、そうした状況に鑑みビブリオメトリックスと称される論文引用関係統計分析の全般について解説し、いわゆる「ランキング」の問題を計算実例に即して検討し、新たな指標の可能性を提示する。すなわち、機関別の h-index の初試算、大学別の引用力評価のための「引用単価」、大学間競争評価のための「大学間競争マップ」、雑誌間競争評価のための「対向係数」等であり、これにより、各々の立場、観点からの主体的・客観的「自己評価」の必要性を提示する。

根岸 正光

(国立情報学研究所教授・SPARC Japan 運営委員長)

1945年生。東京大学経済学部卒、同大学院経済学研究科経営学博士課程修了。東京大学助手、同助教授を経て、学術情報センター教授（データベース研究部門）等。現在、国立情報学研究所、総合研究大学院大学教授。2008年10月、第37回情報化月間・情報化促進貢献・経済産業大臣表彰。著書：「研究評価—研究者・研究機関・大学におけるガイドライン」（丸善、2001）他。

インパクト・ファクター（IF）とは

インパクト・ファクター(IF: Impact Factor、衝撃係数)とは特定の一年間において、その雑誌に掲載された論文が、平均何回引用されているかを示す尺度で、毎年トムソン・ロイター社（以下通称のISI）発行の Journal Citation Reports (JCR)で毎年公表されます[1]。これは、学術雑誌の評価指標として有用であり、よく参照されます。本来、特定の

研究分野における雑誌の影響度を測る指標として利用すべきもので、個々の掲載論文の質を直接示す指標ではありません。しかし実際には、購読雑誌の選定という本来の用法ではなく、研究者評価のために個別論文の評価指標として利用されることが多いようです。

ここで IF の上位の雑誌を調べると、Cancer Journal for Clinicians が 69.03 で第1位でした。IF については、レビュー論文の比率に影響されるこ

とを考慮する必要があります。Cancer Journal for Clinicians の場合、レビューの比率は33%で大きな数字ではありませんが、Annual Review of Immunology は100%レビュー論文のレビュー誌です。もっとも Physical Review は Review といいながらレビュー誌ではないというように注意が必要です

日本誌の IF について調べると、Journal of Photochemistry が 5.73 で日本誌の中では第1位ですが、全体では 273 位です。次が Plant and Cell Physiology で全体の 691 位というようにあまり高い順位ではありません。IF の間違った使い方については、考案者のガーフィールド先生はじめ文部科学省等の指針などでも注意されていますが、それほど本来の目的とは違った使われ方が多くなされているということでしょう。

IF 以外の指標

最近、IF 以外の指標も少しずつ出てきています。一つ目はエルゼビア社の Scopus を使った SCImago です[2]。これは、Google のページ・ランク式の評価法で雑誌をランキングするもので、エルゼビアとスペインのグラナダ大学が行なっています。二つ目は Publish or Perish (PoP) という名前の指標で、オーストラリアのメルボルン大学の別会社を作っているらしく、Google Scholar を自動検索して、オンラインで集計するような仕掛けのようです[3]。三つ目は Journal Usage Factors で、これは、UKSG という英国の学術雑誌編集者団体が提唱しているもので、電子ジャーナルでのダウンロード数によって評価しようというものです[4]。

論文引用統計による評価指標の注意点

そもそも論文引用関連の統計による評価指標において、どんな諸元があるかということ、まず評価の対象の単位が論文レベル、個人レベル、機関レベル、国レベル、雑誌レベルと、いろいろなくくり方があります。また、統計の元になる原デー

タ（データベース）についても、論文数（文献抄録型データベース）、引用数（引用索引データベース）、ダウンロード数（電子ジャーナル）などがあります。

ここで大きな問題は、原データのデータベースの採録範囲がどのようになっているかということです。インパクト・ファクターの原データベースである ISI の SCI (Science Citation Index) は、雑誌を厳選した格好のものであるので、権威のある雑誌の論文相互間での引用統計であるとみなされます。Scopus の方は採録範囲は幅広く、それを使ったのが SRJ です。Google Scholar はウェブ上で公開されている論文をかき集めているため、さらに範囲が広くなり、その上で計算しているのが PoP です。このように、統計の元になっているベースが違うということは十分に理解しておく必要があります。また、統計の観察期間も非常に大事で、論文の寿命、賞味期限が分野によって相当違うという点を考えて検討する必要があります。

ビブリオメトリックスとは

ビブリオメトリックス (Bibliometrics、計量書誌学) は主に論文の引用関係を統計分析するような研究分野で、わが国では図書館情報学系の専売特許のようですが、ヨーロッパ辺りでは主に科学政策論、科学社会学系などで研究され、政策への適用や科学コミュニケーションへの応用を意図した研究が行われています。オランダのライデン大学の CWTS (Centre for Science and Technology Studies) が有名で、ここは「ライデン・ランキング」といわれる欧州の大学ランキングを発表しています[5]。これは論文あたりの引用数を加工した「クラウン指標」によるランキングで、インターネットで公開されています。

さまざまな評価指標

世間では株式投資のために企業評価のさまざまな指標が考案されていますが、大学ランキング

の方でも同様に各種の指標があちこちで考案されています。私はいわば「根岸式総合引用度偏差値」(後出)を案出して、これによるランキングを朝日新聞社の「大学ランキング」誌に毎年掲載しています[6]。終わりにのべるとおり、こうした指標は、そもそも各自、各機関がその用途、目的に即して自ら開発し、使いこなすことが大切であるというのが、本講での私の結論です。その上で、インパクト・ファクターを用いるというのもひとつの選択といえます。

こうした各種指標については、Karolinska Institutet が便利な冊子をインターネットで公開しています[7]。Karolinska Institutet はノーベル医学・生理学賞を出しているスウェーデンの医学校ですが、近年、自らの研究を細かに評価し研究戦略を策定するために、Bibliometric Analysis を導入しているとのこと、そのマニュアルも公開されています。上記冊子にはたくさんの種類のインディケーターが掲げてありますが、最近よく使われるのは h-index で、これについては後程実例をあげます。

ところに SCI とはどんなものかといった問い合わせがありました。当時一般の認識はこの程度であったのですが、その後、世の中で評価が多く使われるようになり、SCI だのインパクト・ファクターだのがことあるごとに話題になり、この十数年で様変わりしました。

私は 2002 年以來、ISI の National Citation Report for Japan (NCR) を使って大学別論文数や引用度を調査した結果を、この「大学ランキング」誌に公表していますが、今年は昨今の世知がらい情勢をにらんで、引用力による予算配分シミュレーションや引用の単価計算を追加しました[6, 8]。

英文機関名の名寄せの問題

さて、実際に統計を作成する上では、年金問題で近頃話題の「名寄せ」という厄介な問題があります。大学、会社などの機関単位で名寄せ集計しようというとき、その書き方にさまざまなバリエーションがあるため、大変な作業になります。英文論文に著者が自分の所属機関を書くときに、独自に英語名を発明してしまうこともあるようで、例えば東京農工大学は標準的略記法では Tokyo Univ Agr & Techno のようですが、このほかに Tokyo Nohkoh Univ 等々、様々に書く先生がおられるため、これらを逐一実態調査をしながら名寄せをしなければなりません。また、表記上、東京農業大学との混同もしばしば起こり、郵便番号や住所などを使って仕分けをしています。静岡大学と静岡県立大学は University of Shizuoka と Shizuoka University なのですが、カレーライスとライスカレーのように間違えやすくなっています。

さらに厄介なのは、独立行政法人で、たとえば物質・材料研究機構は、以前の無機材質研究所と金属材料技術研究所が一緒になったもので、この場合単純です。一方、農業・食品産業技術総合研究機構は、英文略称 NARO で知られています。この NARO というのはここしばらく変わっていませんが、本名の方はいろいろ変わってきており、この間、農業試験場等、各種の機関を何回かに分

The image shows the cover of the 2009 University Ranking magazine. The main title is '2009 大学ランキング' (2009 University Ranking) with a subtitle '737校完全ガイド' (737 Schools Complete Guide). The cover includes a table of university rankings, a list of featured universities, and various articles. The magazine is published by Asahi Shimbun.

日本の大学ランキング

わが国の大学ランキングとしては、先にのべた、朝日新聞が 1994 年に創刊した年刊の「大学ランキング」がありますが、その創刊のときに、私の

けて統合し傘下に収めています。ごく最近では、農業者大学校を吸収して、正式英文名が National Agriculture and Food Research Organization となりました。ちなみに農業者大学校というのは英語で National Farmers Academy というようです。

ともかく 10 年分の統計をとるとなると、昔の名前を調べ、その様々な英文表記を略記法も含め、逐一点検しながら現在の名称に名寄せしなければなりません。分量的にいうと、10 年分の論文 86 万件に対して、著者所属機関レコードは 234 万件という膨大な数に上っています。

引用数の指標化

引用数の指標化は大きな問題です。既にふれた朝日新聞の「大学ランキング」では、分野間の引用数の差を平準化することを前から試みていますが、分野別の引用度指数をそのまま加重平均すると、医科系単科大学の指数が高く出る傾向にあります。これは分布の形状に問題があると考え、引用度の偏差値に換算してから加重平均して「総合ランキング」を出すようにしました。これで機関別の総合指数はほぼ妥当なものになると言えるでしょう。統計の期間は以前は 10 年間で行ってききましたが、2007 年から 5 年間に変えました。5 年では短か過ぎるという意見も聞いていますが。

根岸式引用度指数

分野別引用度指数：論文当たり引用数の分野平均=100で指数化
総合引用度指数：
分野別引用度指数を当該機関の分野別論文数で加重平均

医科単科大の総合指数が高く出がち。
分布形状分野間変異？→「引用度偏差値」（2007年版）
分野別・機関別の論文当たり平均引用数の標準偏差計算
平均=100、標準偏差=30の偏差値に変換

統計期間の短縮 10年→5年（2007年版）
短期的成果、評価の短「気」化に対応
弊害：

- 評価が安定するまでの期間
- 引用がほぼ飽和するまでの期間
- やはり10年程度は必要か（論文寿命）

実際の例ですが、この左頁が総合ランキングで、右側が分野別のランキングです。総合ランキングの元である分野別指標は、すでにのべたような「引用度偏差値」により指数化しています。これを見ると、物理分野で指数が一番高いのは奈良女子大ですが、物理の先生に聞くと、これはまったく意外ということでした。実際には、高エネルギー物理学研究機構等々と共同研究が行われ、その結果として引用度の高い共著論文が出てきているというのが大きな要因です。

The image shows a screenshot of the Asahi Shimbun University Ranking 2009 Edition. It features two main tables side-by-side. The left table is titled 'ISI+論文引用' (ISI+ Paper Citations) and lists various universities with their citation counts. The right table is titled '分野別 (国内2002~2006年)' (Field-specific (Domestic 2002-2006)) and shows citation counts broken down by academic field for each university. A yellow highlight is present on the top right of the page, indicating the source and year of the ranking.

通常のランキングだと、東大、京大等々という常識的な並び方になりがちですが、このような方式、切り口でランキングすると、普段目立たないような大学も上位に現れます。これは実際にその大学にそういう側面があるから現れてくるのであって、分野別ランキングではなるべくこうした新顔が出るように工夫しているわけです。

引用統計の注意点

引用統計の注意点についていうと、とにかく引用は累積効果の産物です。そこで採録雑誌の範囲、統計期間、分野間差異などに注意が必要です。また、共著論文を案分して計算する試みもしており、これについては後で述べます。

ところで、SPARC を始めた頃の話ですが、調査の結果、日本の研究者の論文の80%が海外雑誌に出ているということが分かり、これを海外「流出」だと言ったところ、一部の研究者からこれは海外「進出」であって結構なことだと怒られたことがあります[9]。流出と進出は価値観の問題ですが、ともかく一定の「自給率」は必要であろうと考え、その意味でSPARC Japan の事業は有意義だと思っています。9月15日のリーマン・ショック以降、米の自給率はゼロでいいではないか、もう高い米作りは止めて、その代わりにアメリカに車を買ってもらえば良いといった勇ましい自由貿易説は大分様変わりし、ある程度の自給率は必要という論調が強くなっている感じがしますが、これと似たような状況でしょう。

採録範囲の問題

採録範囲の問題ですが、この7月にISIでは、リジョナル・ジャーナルと称して700誌程を追加採録するという発表がありました。リジョナル・ジャーナルは、インターナショナルではないけれども、地域的に有意な貢献を評価したもので、その中で日本の雑誌も30誌ほどが追加になり、SPARC Japan が支援している雑誌も幾つか入っています。

これについては、ISIの製品を買ってくれば、そのジャーナルは入れていこうというビジネスとしての話も当然あるでしょう。そういう意味で、もっとISIの製品を買えば、さらに日本の雑誌も入るのではないかという感じもします。ISIのデータベースは民間会社がビジネスとしてやっているのですが、昔の名前である Institute for Scientific Information からの連想でしょうが、研究評価のための公的な研究機関がやっているものと誤解している人がいます。ともあれ、日本でもISIの製品をよく使うようになってきたということの一つの表れが、今回、日本誌が追加されたということに出てきているのではないかと、個人的に解釈しています。

研究分野間差異の問題

研究分野間で引用数が相当に違うということについては、生物、医学系では、論文の末尾に書かれる引用文献リストが特に長い、従って論文当たりの引用数が高くなるといった現象で、実際に統計を取ってみると、このようになります。

論文当たり(被)引用数の分野間差異
2002-2006年の日本著者論文に対する同期間の引用
トムソン National Citation Report for Japan 1997-2006より統計

分野別	論文数	引用数	論文当たり引用数
免疫学	4,984	65,821	13.21
分子生物学・遺伝学	10,696	134,955	12.62
宇宙科学	3,362	32,277	9.60
生物学・生化学	31,238	232,717	7.45
学祭研究	3,915	28,116	7.18
神経科学	13,975	95,269	6.82
微生物学	8,614	51,930	6.03
化学	54,204	312,795	5.77
臨床医学	80,017	428,015	5.35
薬学	10,225	47,871	4.68
物理学	73,416	317,962	4.33
地球科学	8,517	36,238	4.25
生態・環境学	4,457	18,155	4.07
動植物学	17,862	67,463	3.78
材料科学	21,422	73,375	3.43
農学	8,601	25,975	3.02
精神医学・心理学	2,118	6,220	2.94
工学	36,574	87,043	2.38
人文社会科学	3,374	4,652	1.38
数学	6,038	8,045	1.33
コンピュータ科学	7,833	6,021	0.77
全体	372,473	1,895,402	5.09

24

この数字だけ見ると、わが国では免疫学が論文当たりの引用数が13.21と一番高くても発達しており、コンピュータ科学は0.77で非常に遅れた分野と判断されそうです。その結果、コンピュータ科学はやめて、免疫学の研究にもっと力を入れていこうというような議論が実際に行われがちなので注意が必要です。

総合科学技術会議での年次評価

国も国立機関の研究評価を定常的にする必要があるので、2004年から総合科学技術会議、内閣府による国立大学法人、研究開発型独立法人の年度評価を始めました。この中にISIデータベースでの引用統計を入れたいという話がこのとき私のところにきました。年次評価だからといって、単に昨年出た論文に対する昨年の引用数だけ見てもこれは乱暴で無意味です。そこで、過去3年間の論文に対する昨年の引用数、2002年から2004年の論文に対する2004年の引用数を統計するというように、引用する年度を限定する方式で

年度評価にすることにしました。このように過去3年間の論文を対象にすることでスタートしましたが、やはりこれでは短すぎると考えられ、翌年から過去10年間の論文に対する当該年度の引用数を統計するというかたちで今年まで続けています。

10月31日の総合科学技術会議に昨年度分の年度評価資料が提出され公表されています[10]。国立大学法人、大学共同利用機関法人、国立高専機構、研究開発型独法、それから主要公私立大学についての調査結果がまとめられていますので、いくつか示します。

国立大学法人等の科学技術関係活動に関する調査結果(平成19事業年度) 平成20年10月31日 内閣府(科学技術政策・イノベーション担当)

別紙1 ISI 論文数(2007年)、引用数(2007年)

Table with 6 main sections: Astronomy (天文学), Chemistry (化学), Computer Science (コンピュータ科学), Earth Science (地球科学), Engineering (工学), and Mathematics (数学). Each section contains a table of institutions, 2007 publications, and 2007 citations.

独法の統計では、すでにのべたように名寄せが非常に大変で、とくに産業技術総合研究所は旧工業技術院相当の大所帯であるので厄介です。その一方で、電子航法研究所や大学では東京外国語大学など、年間数件しかISIの論文のないところは、漏れのないよう、その分よほど綿密に調査しなければなりません。小さいところほど手間がかかるという側面もあるのですが、誤差を小さくするという意味からやむを得ないでしょう。

実際の例をみてみると、独立行政法人については、2007年の刊行論文数などはこのようになっています。ここには内閣府の方で研究者一人当たり割りに算出した数字も出ています。



研究開発型の独立行政法人ということで、大学等と並べてどのぐらいの順位にあるかというデータも出されています。これによると工学、生態、環境学の引用数では産総研がトップとなっています。その他、この資料には各大学の分野別の論文数、引用数なども出ていて詳しい内容はウェブで確認することができます。

Table titled '独立行政法人の研究開発活動について' showing citation data for various research institutes. It includes columns for field, institute name, and citation counts. A note explains the methodology for counting citations.

機関別 h-index の計算

次に触れるのは h-index です[11]。これはビブリオメトリクス業界で昨今流行っている指標で、物理学の先生が発明した勘定の仕方です。論文を

引用数の高いものから順に並べて、だんだんゼロに近いものまで並べる。それを原点から 45 度線で当たったところの数をもってインデックスにしようというものです。これは論文の質と量を一つの数字で表すことができるため便利です。大学ランキングは、私が行っているもののように、やはり論文の数と、論文当たりの引用数あるいはそれを指数化したものとの両方を見比べて評価するべきだと思っていますが、h-index はこの両方を一つの数字で見ることができるというのがミソでしょう。

次の表は、実際に国公立大、独法の h-index を計算してみた結果です。10 年間で東大が 214、阪大が 193、京大が 191、JST が 175、理研が 148 というような数字が出ています。

国公立大・独法の h-index を初試算 1998-2007年・全分野
(原データ: Thomson, "NCR-J 1998-2007" 引用度はこの期間の引用数/論文数)

SEQ	機関名	10年間論文数	引用度	H指数	SEQ	機関名	10年間論文数	引用度	H指数
1	東京大学	65,081	12.9	214	26	大阪市立大学	7,986	10.0	84
2	大阪大学	40,546	12.7	193	27	情報・システム研究機構	2,637	15.6	84
3	京大	46,916	12.3	191	28	横浜国立大学	4,251	12.3	82
4	科学技術振興機構(独)	18,944	17.4	175	29	岡山大学	12,264	8.4	81
5	理化学研究所(独)	16,142	14.9	148	30	首都大学東京	5,866	10.8	81
6	東北大学	38,704	9.8	145	31	東京理科大学	7,254	8.4	79
7	名古屋大学	25,568	10.6	128	32	信州大学	6,980	8.8	77
8	九州大学	26,526	9.3	120	33	昭和大学	3,647	12.2	77
9	筑波大学	16,642	9.7	114	34	久留米大学	2,945	13.4	77
10	自然科学研究機構	10,240	13.0	114	35	物質・材料研究機構(独)	9,849	7.6	75
11	産業技術総合研究所(独)	25,973	8.8	112	36	群馬大学	6,548	8.9	75
12	東京工業大学	22,727	9.2	112	37	長崎大学	6,792	9.1	74
13	北海道大学	25,742	8.8	107	38	徳島大学	6,545	9.2	74
14	慶應義塾大学	11,750	10.0	103	39	名古屋市立大学	4,118	11.3	74
15	千葉大学	11,535	10.3	103	40	東京女子医科大学	4,114	10.6	74
16	神戸大学	10,518	9.7	98	41	早稲田大学	7,066	7.5	72
17	熊本大学	7,213	11.9	98	42	鹿児島大学	5,586	8.2	72
18	東京医科歯科大学	7,086	13.3	98	43	三重大学	4,945	9.2	71
19	順天堂大学	4,094	14.7	97	44	総合研究大学院大学	2,682	13.1	71
20	新潟大学	5,323	10.6	95	45	岐阜大学	6,561	8.9	69
21	広島大学	14,987	8.5	92	46	愛媛大学	5,008	9.1	69
22	奈良先端科学技術大学	3,156	15.8	91	47	自治医科大学	3,596	10.8	69
23	金沢大学	8,413	10.4	89	48	関西医科大学	2,226	13.2	68
24	高エネルギー加速器研	5,441	12.5	89	49	兵庫医科大学	2,224	12.9	68
25	東海大学	5,034	11.6	86	50	札幌医科大学	3,291	11.1	67

ここで材料分野だけの結果を出すと、東北が 54 で 1 番、次が産総研、物材研、九大、阪大という順になっています。こうした h-index の値は、大体その分野の研究者の実感にマッチした数字が出てきているように思われます。研究規模が大きく、研究レベルも高い順に出てくるため、そういう意味では妥当性があるように見えますが、基本的には論文数に比例する側面が非常に強く、機関の規模に準じた「平凡な」ランキングになっています。h-index は元々個人評価を念頭に案出された指標なので、規模が非常に異なる機関間にその

まま適応するのは無理があり、何らかの標準化が必要だと思います。

国公立大・独法の h-index を初試算 1998-2007年・材料分野
(引用度はこの期間の引用数/論文数)

SEQ	機関名	10年間論文数	引用度	H指数	SEQ	機関名	10年間論文数	引用度	H指数
1	東北大学	4,312	6.8	54	26	東京農工大学	334	5.9	22
2	産業技術総合研究所(独)	3,810	7.2	51	27	京都工芸繊維大学	518	4.6	21
3	物質・材料研究機構(独)	2,871	6.6	46	28	長岡技術科学大学	463	4.9	21
4	九州大学	1,325	7.8	45	29	理化学研究所(独)	276	6.5	21
5	大阪大学	2,594	6.8	44	30	神戸大学	223	7.3	21
6	東京大学	2,515	6.2	43	31	大阪市立大学	133	9.0	21
7	科学技術振興機構(独)JST	1,004	10.0	41	32	熊本大学	336	5.6	20
8	京都大学	2,184	6.1	39	33	首都大学東京	302	6.2	20
9	東京工業大学	2,161	6.7	38	34	千葉大学	236	5.4	19
10	名古屋大学	1,360	5.7	33	35	日本原子力研究開発機構(独)	396	3.8	18
11	大阪府立大学	756	8.0	32	36	群馬大学	202	5.7	18
12	早稲田大学	491	7.9	28	37	宇宙航空研究開発機構(独)	300	4.4	17
13	北海道大学	1,205	5.2	27	38	岡山大学	280	4.5	17
14	兵庫県立大学	421	7.8	27	39	福山大学	264	4.3	17
15	慶應義塾大学	479	8.4	26	40	九州工業大学	322	3.8	16
16	信州大学	387	7.2	26	41	岐阜大学	311	5.0	16
17	福岡教育大学	46	44.3	25	42	電気通信大学	247	4.9	16
18	国立高等専門学校機構(独)	661	4.6	24	43	山形大学	244	5.1	16
19	豊橋技術科学大学	495	6.0	24	44	徳島大学	178	5.6	16
20	名古屋工業大学	792	4.6	23	45	新潟大学	166	6.1	16
21	筑波大学	491	5.3	23	46	山梨大学	152	8.2	16
22	広島大学	464	6.3	23	47	佐賀大学	144	6.6	16
23	東京理科大学	489	5.5	22	48	自然科学研究機構	132	7.5	16
24	静岡大学	485	5.6	22	49	東海大学	261	4.1	15
25	横浜国立大学	366	6.4	22	50	鹿児島大学	212	5.2	15

ところで h-index は、アイデアは非常にシンプルだが手順的な計算過程を含むので、機関別の h-index の計算となると、大量のデータを多数の機関について処理する必要があり、非常に厄介です。いろいろ工夫した後、効率的な計算方法を考え出した結果得られたのが先ほどの数値です。実際の結果をみると、確かに研究者達の総合的な評価、「実感」に合致するというのはよい点でしょうが、結局規模の影響が大きいということで、何らかの補正が必要です。やはり論文数と引用度指標の両方を見比べるという、複眼的思考が重要ではないかと考えています。

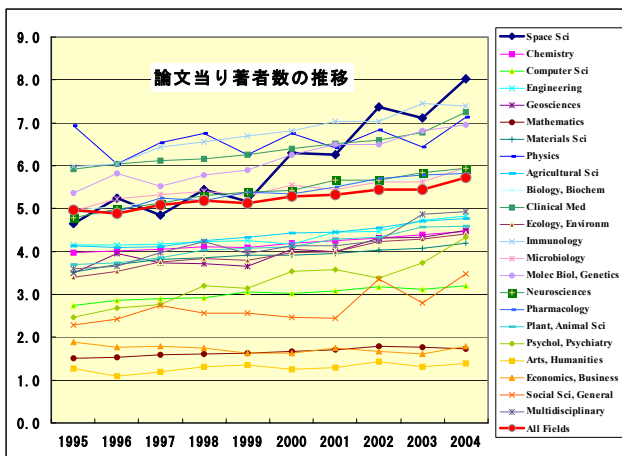
納税者評価

さて「大学ランキング」の今年版に入れた新しい統計、「引用力による予算配分と引用単価」の話です。研究評価ではピア・レビューという言葉がよく使われますが、つまりは同業者評価なので、どうせ素人には分からないだろうとごまかす可能性もあります。研究者が税金を使って真面目に研究に取り組んでいるのかどうか、これを検証する意味で、納税者の視点からの評価という考え方が重要になってきていると思われます。

2006年度の国立大学の交付金、私学補助金、科研費を足すと、1兆5670億円となっていて、この春先に話題になった道路特定財源5.4兆円の中のガソリン税暫定増額分がちょうど1.4兆円でしたので、同程度の税金が国立大、私立大に使われているということになります。1兆円という大きな数字ですが、これだけ大きな全体予算の中では、むしろ割と少ない額ではないかと思えます。産業利益にすぐに貢献するような有用な研究成果が大学方面から出てこないのはけしからん、税金のムダづかいという納税者の声もあるようですが、ガソリン税の暫定増額分程度しか出してないわけで、言う方がおかしいような気も致しますが、ともあれ資金効率的観点での計算を試みた次第です。

共著者数、共著機関数の増加

そこで先ず、これは論文数当りの著者数の趨勢を分野別に示したグラフです。論文あたりの著者数がどんどん増えてきていて、特にSpace Scienceではここ10年非常に伸びています。共著機関の数で勘定しても同様の傾向です。



共著者数の多い実例を見ると、例えば日本循環器学会の英文誌、Circulation Journal, Vol.68, p.860-867 (2004)にある論文ですが、2459名の共著者による論文です。コレステロール、高血圧に関する論文で、全国の病院の先生に声を掛けてデー

タを集めたため、それに協力した先生2459名の名前がずっと並んだ共著論文になっています。他の例では、PHENIX detectorという加速器関連のビッグサイエンスの総括論文だと思いますが、588名の著者の名前だけで実に3ページ、所属機関だけで1ページを丸々使って、5ページ目からようやく本文が始まるというような論文も世の中にはあるのです。

共著論文の論文数、引用数の案分計算

このように共著者数が増えると、どうも評価対策の「水増し」の可能性があるのでないかと考えられてしまいます。そこで論文数、引用数を共著者数、共著機関数に応じて案分して補正するとどうなるか、この計算もしてみました。実際、単純集計だと、共著の影響で先ほどの奈良女子大の話と同様に、国学院大が物理で上位にランクされることもありました。また、東京外語大は論文が非常に少ないのですが、その中に医学系論文が数件ありました。調べると保健センターの先生の論文で、たぶん東大医学部の出身ではないかと思えますが、東大医学部との共著論文でした。

このような具合なので、共著論文は案分計算するのがある意味で正しい方法だと思います。本来は著者数に応じて案分するのが妥当ですが、ISIデータベースのデータの構成上それができないため、機関の数で案分しました。以前にこれで「大学ランキング2007年版」を出したことがありますが、論文数が平均して6掛け程度に減ってしまうので評判が良くなかったため、これは元に戻しました。もっとも「引用力」によって交付金の配分をするという、今回の資金効率計算の場合は、案分論文数、案分引用数で計算するのが適当と考え、引用力はこの方式で計算しています。

「引用力」による予算配分と引用の単価

近ごろは、国の財政状況から予算がシビアになってきており、私学方面からはイコール・フッティングと称して、国立大への交付金は多すぎるの

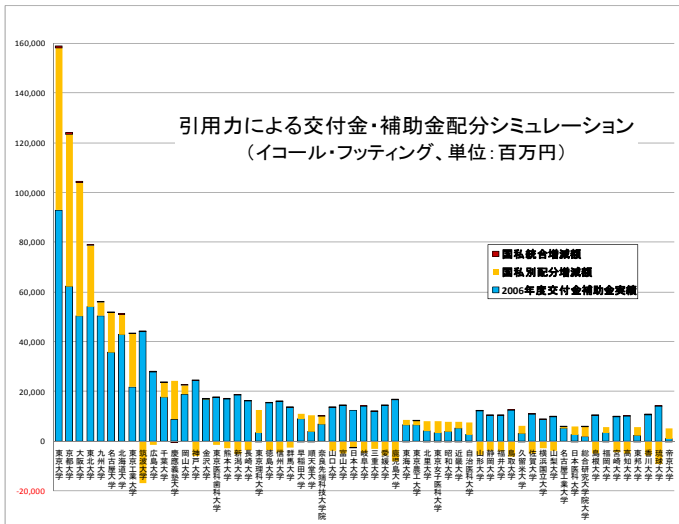
ではないか、これを私立大学補助金の方へ回す方がよいという議論があるようです。この点を考慮した上で実際に引用力による交付金・補助金配分のシミュレーションをしました。計算してみると、むしろ私学への配分額を 2.2%減らして、これを国立の方へ振り替えるという興味深い結果になりました。つまり引用力に比例した予算配分を考えると、すでにイコール・フッティングが実現されているということになります。しかしこれを大学別で見ると、東大など大規模大学にはさらに予算を増額すべしという恐るべき？結果になっています。

引用の「単価」

案分被引用数(2002-2006年)1件あたりの交付金・補助金+科研費取得額

(2006年度国立大運営費交付金、私立大経常費補助金と取得科学研究費の合計値により算出)
(全87国立大と案分被引用数1,000件以上の私立大引務について試算)

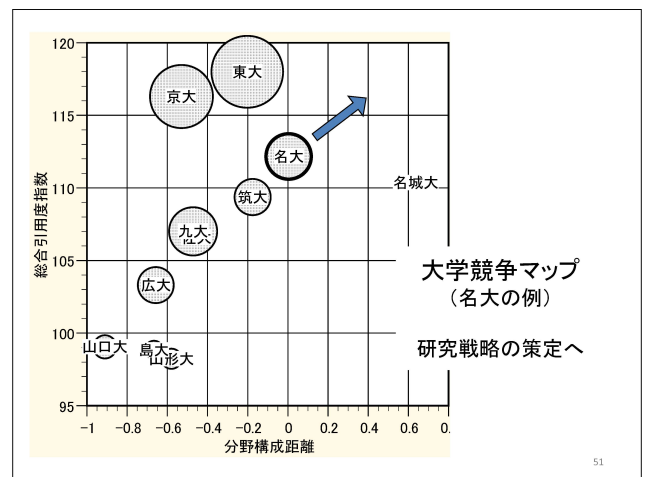
順位	大学	引用当り交付金補助金+科研費(千円)	2006年度交付金+補助金+科研費(千円)	案分被引用数	順位	大学	引用当り交付金補助金+科研費(千円)	2006年度交付金+補助金+科研費(千円)	案分被引用数
1	京都薬科大学	184	673	3651.5	26	北里大学	784	4,495	5730.5
2	神戸薬科大学	227	303	1336.7	27	東京医科大学	793	2,208	2783.3
3	星薬科大学	231	510	2210.1	28	徳島文理大学	806	1,137	1411.3
4	東北薬科大学	284	294	1037.0	28	大阪医科大学	806	1,492	1851.8
5	帯広大学	284	1,081	3729.7	30	豊田県立大学	823	2,045	2485.4
6	東京薬科大学	300	1,014	3375.0	31	新潟県立大学	829	5,107	9062.0
7	明治薬科大学	326	433	1328.1	32	東京工業大学	834	26,431	31699.3
8	豊田工業大学	429	571	1329.9	33	名城大学	848	1,737	2048.8
9	東京理科大学	436	3,934	9023.6	34	岡山理科大学	870	1,280	1471.4
10	総合研究大学院大学	455	1,862	4310.4	35	筑波医科大学	881	2,411	2739.1
11	川崎医科大学	490	897	1830.7	36	北海道医療大学	885	1,117	1262.5
12	自治医科大学	516	2,830	5485.7	37	福岡大学	897	3,632	4048.3
13	兵庫医科大学	541	1,682	3108.7	38	神奈川大学	921	1,649	1790.2
14	順天堂大学	576	4,318	7501.5	39	愛知医科大学	970	1,335	1376.9
15	東邦大学	592	2,365	3992.2	39	近畿大学	970	5,423	5588.3
16	金沢医科大学	596	1,008	1736.4	41	東京大学	972	112,829	116061.0
17	東京女子医科大学	618	3,532	5715.6	42	埼玉医科大学	989	1,807	1809.0
18	慶應義塾大学	636	10,999	17284.1	43	奈良先端科技大	1,042	7,724	7412.9
19	関西医科大学	639	2,310	3617.8	44	東京歯科大学	1,053	1,201	1141.0
20	日本医科大学	646	2,788	4317.7	45	千葉大学	1,109	19,392	17489.5
21	獨協医科大学	660	2,127	3220.7	45	東北大学	1,109	64,052	57732.6
22	杏林大学	717	1,810	2523.4	47	東海大学	1,118	6,897	6188.0
23	久留米大学	722	3,281	4546.7	48	名古屋大学	1,120	42,555	38006.0
24	昭和大学	724	4,109	5674.6	49	聖マリアンナ医科大学	1,188	2,344	1973.3
25	大阪大学	773	58,997	76285.7	50	東京農工大学	1,224	7,484	6114.6



「ランキング」の本質の問題点と「大学競争マップ」

ランキングとは一次元に並べて配列することなので、それ自体、相当に無理があるものです。もう少し多次元的な中で考える必要があるのではないのでしょうか。その点を踏まえて、以前に研究分野の類似性、すなわち論文分野別構成比による主成分分析をしたことがあります。それによって、機関相互間の「距離」を算定し、論文数や引用度偏差値の倍率等々を勘案して、球型散布図でこれを一度みられるようなグラフ、すなわち「大学競争マップ」を作ってみました。

次に、引用の単価の計算をしました。1 引用を獲得するのにどの程度の予算を使ったかという表ですが、これをみると京都薬科大学が一番安上がりで 18.4 万円なのに比べて、東京大学は 100 万円弱かかっています。この単価計算では、さすがに私立大学の方が資金効率良く引用を獲得していることが分かります。だからといって、京都薬科大学や神戸薬科大学の方に予算を重点的に配分していけば、引用数がどんどん上がるかというと、そう簡単ではないことはいうまでもないでしょう[12]。



これは名古屋大学について計算をしたもので、名古屋大学を中心にして周りにいろいろな大学があるという図式が描けます。名大が左上の東大の方へ向かって突き進むのか、右上の空いている方向へ行くのがいいか、研究戦略の今後の持っていくよという事で、面白いのではないかと思います。その後、この研究はあまり進めてませんが、今後やってみたくと思っています。

雑誌間の競争関係を計測する「対向係数」

今のは大学・機関単位の話ですが、雑誌単位の競争関係についても JCR のデータを使って以前検討したことがあります[13]。これは、電子情報通信学会の雑誌の現状評価というテーマで、この学会誌に掲載したもので、詳しくはこの論文を見ていただければ分かりますが、要するにある雑誌が引用したり、または引用される雑誌はどういうものであるか、その競争関係を定量的指標で表そうという試みです。

この場合、引用比を \log_2 にすると、ちょうど直感的に見やすい数字になることが分かり、この指標を「対向係数」(RF: Reciprocity Factor)と命名しました。

対向係数 (Reciprocity Factor: RF)

- 引用比 = 被引用数 / 引用数
- 『対向係数』 = $\log_2(\text{引用比})$
- 引用比1(互角)のとき対向係数は0
- 対数の底を2。引用比2または1/2のとき対向係数は±1となり、双方の格差が2倍以内のときには1以内。
- 「優劣」を直感的に把握
- IEICE Trans. Communからみて、IEEE COMMUN LETT: -1.83やIEEE T VEH TECHNOL: -2.01で、相対的に近い。それでも-1を超えている、すなわち引用・被引用に2倍以上の差。その克服には相当の対策が必要。(次掲、表2)

電子情報通信学会の Trans. Commun を中心にした対向係数をみると、このような結果になります。競争相手として、一番近いと思われるのは、IEEE の Commun Lett ですが、それでも-1.83 で、倍近

く突き放されており、これに追いつくには相当な努力が必要だというのが分かります。

表2 IEICE Trans. Commun.の類縁誌とその引用統計・対向係数 (ISI JCR on CD-ROM, Science Edition 2002より集計。対向係数: $\log_2(\text{引用比})$)

引用・被引用誌	発行国	IF	論文数	引用数(A)	被引用数(B)	合計(A+B)	引用比(B/A)	対向係数
1 IEICE T COMMUN	JPN	0.487	319	234	234	468	1.00	0.00
2 IEEE J SEL AREA COMM	USA	2.316	149	175	16	191	0.09	-3.45
3 IEEE T COMMUN	USA	1.562	223	145	9	154	0.06	-4.01
4 IEEE COMMUN MAG	USA	3.165	178	135	2	137	0.01	-6.08
5 IEEE T VEH TECHNOL	USA	1.220	142	101	25	126	0.25	-2.01
6 IEEE ACM T NETWORK	USA	2.408	65	109	4	113	0.04	-4.77
7 ELECTRON LETT	ENG	1.072	1119	96	12	108	0.13	-3.00
8 IEICE T FUND ELECTR	JPN	0.336	317	33	65	98	1.97	0.98
9 IEEE COMMUN LETT	USA	1.220	177	32	9	41	0.28	-1.83
10 IEEE T SIGNAL PROCES	USA	1.159	280	41	41	0.00	-99.99	
ELECTRON COMM JPN 1	JPN	0.018	99		30	30		99.99
IEICE T ELECTRON	JPN	0.571	302	13	16	29	1.23	0.30
IEICE T INF SYST	JPN	0.148	193	1	9	10	9.00	3.17
JCR全体・全体対向係数				3921	864	4785	0.22	-2.18

この計算法を JCR データベース全体に対してやってみると、これを「全体対向係数」と命名しましたが、例えば Nature は 2.87、Science 2.85 などで、インパクト・ファクターと比較すると面白い結果が出そうなのでさらに検討したいと思っています。

「全体対向係数」とIF

- JCR全体との対向係数: JCR全雑誌との間での引用・被引用の集計値から求めた対向係数。
- NATUREの対向係数: 類縁誌に対して+の係数にならない。BIOL CHEMに対して3.90、P NATL ACAD SCI USAでは3.62、SCIENCEには0.02、PHYS REV Bに対しては3.89。SCIENCEに対しては0.02、すなわち引用・被引用がほぼ同数ながら辛勝。
- 有名誌の全体対向係数: NATURE 2.87(IFは30.432)、SCIENCE 2.85(同26.682)、P NATL ACAD SCI USA 1.37(同10.700)。

むすび

ランキングは、どこかの他人が作ったものを見て、そこで一番になっているからといって喜んだり、成績不振のときには「こんな順位にあるのはおかしい」と言ったりすることが多いのですが、

私にいわせれば、ランキングは本来、他者からの評価をはね返すような自己評価に使うというのが正しい使い方であると思います。自分の一番いいところ、自分の取り柄は自分で探すべきであって、そのためにこの種のデータを使うということです。

従って、客観性、説得性のある評価指標の発見・開発を、各自の立場でやっていく必要があると思います。これは、「後ろ向きの評価対策」から、「攻めの自己アピール」への転換ということになるでしょう。ただしこれにはそれ相応の時間、暇、金がかかる、その覚悟は必要です。どこかでだれかが作ったランキングに振り回されるのではなく、自分の立場、視点でこの種の指標を開発し、ランキングを作っていくという姿勢が研究機関等では必要だと思えます。

さて、もともとわが国では「番付」と称して、ランキングが非常に好まれています。「番付で読む江戸時代」という本は大変面白く、明治20年代までの番付3,294件の番付目録を収録しています[14]。この表紙に「あほうとかしこの番付」というのが出ています。「かしこ」の方の大関を目指すのがまともでしょうが、この際「あほう」の方でもいいから、ともかく大関を目指して、それぞれの立場からランキングに向き合っていくのが、何より大切な姿勢ではないでしょうか。

参考資料

- [1] JCR: Journal Citation Report
<http://www.thomsonscientific.jp/products/jcr/>
<http://www.thomsonscientific.jp/products/jcr/support/faq/#1>
- [2] SCImago Journal Rank (SJR)
<http://www.scimagojr.com/index.php>
- [3] Publish or Perish (PoP)
<http://www.harzing.com/pop.htm>
- [4] Journal Usage Factors (JUF)
<http://www.uksg.org/usagefactors>
- [5] Leiden Ranking, CWST, Leiden University
http://www.cwts.nl/cwts/LR_green_table.html
- [6] 根岸正光「ISI・論文引用ランキング：『引用力』による配分額を初試算」,「大学ランキング2009年版」(朝日新聞社), p.216-225 (2008.5) ISBN978-4-02-274520-0
- [7] Karolinska Institutet, "Bibliometric indicators? Definitions and usage at Karolinska Institutet," 33p. (2007)
http://ki.se/content/1/c6/01/79/31/Bibliometric%20indicators%20-%20definitions_1.0.pdf
- [8] NCR: National Citation Report
http://www.thomsonreuters.com/products_services/scientific/National_Citation_Report
- [9] 根岸正光「研究評価における文献の計量的評価の問題点と研究者の対応」薬学図書館, 49巻3号, p.176-182 (2004)
- [10] 総合科学技術会議 (第77回) 議事次第・配布資料 (2008.10.31)
<http://www8.cao.go.jp/cstp/siryu/haihu77/haihu-si77.html>
- [11] Hirsch, J. E. "An index to quantify an individual's scientific research output." Proceedings of the National Academy of Science of the United States of America. vol.102, no.46, 2005, p.16569-16572.
<http://www.pnas.org/content/102/46/16569.full.pdf>
- [12] 根岸正光「ISI データベースにおける論文の『引用力』による交付金・補助金の大学別配分シミュレーション」情報知識学会誌, Vol.18, No.2, p.131-142 (2008.5)
<http://ci.nii.ac.jp/naid/110006782179/>
http://www.jstage.jst.go.jp/article/jsik/18/2/131/_pdf/-char/ja/
- [13] 根岸正光他「引用度からみた電子情報通信学会発行誌の評価」電子情報通信学会誌, Vol.87, No.9, p.770-775 (2004.9)
- [14] 林英夫他「番付で読む江戸時代」柏書房 (2003)