

ビブリオメトリクスの方法とその応用

Bibliometrics: Methods and Applications

原 田 勝

Masaru Harada

Résumé

In this paper, the author reviews the methods of bibliometrics and its applications.

Analysis of written communication is an essential tool to elucidate the characteristics of science, scientists and scientific literature. Part of this theme has been studied as the statistical bibliography concerned with gross trends of scientific progress, and as reference counting performed by library scientists. Recently, these two areas have been united and named 'bibliometrics.'

Many factors have contributed to the formation of this branch of study; claims for the fundamental theory of library and information science, and the recognition of the importance of systems approach and user study for solving library problems.

However, because the term 'bibliometrics' is used exclusively in the library science field, this subject tends to be restricted to the problems of selection of journals in the library. But the scope of the applications of bibliometrics should not be restricted. The usefulness of the methods has also been recognized by sociologists and historians of science and they provided us with many useful findings.

The main topic of bibliometrics is the Bradford's law of scattering. Although his motivation of the study was to perform complete documentation, it is considered as the useful means to select scientific journals in the library. But the law involves some ambiguities and limitations. The law could have been useful if it were combined with the methods of systems analysis.

The methods of bibliometrics can be applied to classification of scientific papers, elucidation of the characteristics of scientific journals and disciplines, analysis and prediction of scientific progress, selection and retirement of scientific journals in the library, and checking up of defects of bibliographical services.

- I. は じ め に
- II. 図書館学における科学文献の統計的分析研究の必要性
 - 1. 図書館情報学の基礎理論
 - 2. システム分析
 - 3. 利用者研究
- III. 歴史的回顧
 - 1. 統計的書誌学
 - 1) コールとイールズの研究
 - 2) ヒュームの統計的書誌学
 - 3) ロトカの法則
 - 2. 引用文献調査
 - 1) タイトルの分散
 - 2) 時間的広がり
 - 3) 言語の分布
 - 4) 主題の分散
 - 5) 文献の形態
- IV. ブラッドフォードの法則
 - 1. ブラッドフォードによる法則の定式化
 - 1) 実験の背景
 - 2) 法則の定式化
 - 2. 法則の数学的表現
 - 1) 法則の数学的表現
 - 2) 文献の特性を表わすパラメータ
 - 3. 実験方法
 - 1) 各種の実験方法
 - 2) 一次雑誌の引用文献による方法の特長と欠点
 - 3) 実験方法の選択
 - 4. 文献の寿命と成長
 - 5. いくつかの注意
- V. ビブリオメトリクスの応用
 - 1. 理論的応用
 - 1) 文献の分類
 - 2) 科学雑誌・学問領域の特性
 - i) 科学雑誌の特性
 - ii) 学問領域の特性
 - 3) 科学の進歩の分析と予測
 - i) 情報伝播の疫病モデル
 - ii) 科学研究のライフ・サイクル
 - iii) 科学発展のマルコフ・モデル
 - 2. 実応用

- 1) 図書館における雑誌の収集と廃棄
- 2) 書誌の完全性

VI. む す び

I. は じ め に

記録情報の統計的分析は、科学文献の特性の解明と同時に、科学と科学者の特性を明らかにするためには不可欠の手段である。このような研究はこれまで「統計的書誌学」、「引用文献調査」などと呼ばれてかなり多く行なわれてきたが、その研究範囲は限られたものでしかなかった。しかし、近年、科学・科学者・科学文献の特性を明らかにし、有効な図書館活動を行なうためには、記録情報の多面的な分析が必要であることが認識され、この分野に対して「ビブリオメトリクス」(bibliometrics)なる名称がブリチャードにより提案された。彼はこれを次のように定義している: “記録されたコミュニケーションのさまざまな局面を計数し分析することによって、記録されたコミュニケーションおよびある学問領域の発展の特性と流れ(ただし、これが記録されたコミュニケーションによって示されている限りにおいて)の過程を明らかにする(知識の一分野)。”¹⁾

この名称は次第に受け入れられているが、その研究範囲はあいまいで、ビブリオメトリクスの手法を応用することによって何を明らかにすることができるかについて、個別的な研究成果の蓄積はあるが、総合的な観点から書かれたことはなかった。本稿は、ビブリオメトリクスを科学・科学者・科学文献のさまざまな特性を明らかにする目的で行なわれる記録情報の数学的・統計的分析手法の総体と考えて、その諸方法と応用範囲のいくつかを述べる目的で書かれた。

II. 図書館学における科学文献の統計的分析研究の必要性

ビブリオメトリクスに関する論文の数は、1960年代とくにその後半に入ってから、急速に増加する。これは、第一に、巨大科学の特性を明らかにするために、科学組織、科学者、科学文献の統計的分析がプライスによって行なわれ、その手法の重要性が広く認められたことによる。第二に、科学の社会システムを解明するための一方法として、科学社会学者たちにより、科学文献の分析が始められたことによる。第三に、ブラッドフォードの法則の重要性が図書館学研究者の間で再認識されたことによる。

このような事情を背景として、科学文献の統計的研究が一つの重要な研究テーマと考えられるようになり、これに対してビブリオメトリクスなる名称が与えられた。しかし、ビブリオメトリクスが一つの重要な方法として提案されたのは、他分野からの刺激やブラッドフォードの法則の再評価にのみよるのではなく、現在の図書館学研究がこのような手法を必要としているからに他ならない。もちろん次に述べる諸テーマへの関心を、他分野における研究とともに、ビブリオメトリクスの形成を促した諸要因として記述しなすことも可能であるが、ここでは上に述べた観点から図書館学研究の現状を概観してみることにする。

1. 図書館情報学の基礎理論

情報の性質・蓄積・検索・流通・利用などに関する諸問題をより広い視野から研究することを目的として(あるいは、二つの分野の研究目的の類似性により)、図書館学、ドキュメンテーションを総合して生まれた図書館情報学も、コンピュータ技術者の開発した情報の蓄積と検索の技術、図書館学の伝統的な目録と分類の技術という二つの技術を有するのみで、これら二つの前身に共通して欠けていた基礎理論を未だつくりあげていないことは多くの論者によって指摘されている。例えば、シェラは、基礎理論の不在を指摘して、次のように述べている。“図書館学は、未だ、すべての図書館員および情報学者が各々の専門分科に入る前に修得すべき、共通の一般化された学問内容を進歩させたと言えるほど充分には成熟していない。”²⁾

図書館学に基礎理論・哲学が欠けていることは事実であり、その事実を指摘することは容易であるが、重要なことは、その指摘にこたえて、たとえ充分に完成されたものでなくとも、研究の手がかりとなりうる方法を提示し理論を構成しうる要素を提示することによって、基礎理論をつくりあげていくことであろう。しかし、基礎理論の欠如は、図書館学の研究とは何かについての考え方で変えてしまうおそれがある。

1966年から70年までの5年間の英国における図書館学に関する研究のレビューの中で、イェランドは“1965年以前に、大規模で組織的な図書館研究があったとはほとんど言えない”³⁾と述べ、1965年教育科学省に、科学

技術情報局 (Office for Scientific and Technical Information, OSTI) が設けられたことがこの分野の研究の発展にとって重要なステップであったことを強調して、この5年間に行なわれた研究を紹介している。しかし、イエランドの言う“大規模で組織的な研究”とは多くの予算と人員を投入した研究の意味であり、事実、その後が続く記述の主要な部分は、利用者研究、機械化およびシステム分析に関する研究である。もちろんこれらは図書館学の主要な研究テーマとなるべき課題であるが、基礎研究はそれ以上の意義を与えられるべきである。彼のレビューを見ればわかるように、Classification Research Group の活動を除くと、基礎研究の分野で紹介に値する研究はなかった。

同じレビューの「情報学」(Information Science) の項で、ブルックスは、情報プロセスの機械化に関する研究は情報学に何もつけ加えなかったこと、情報学をシャノンの情報理論と関連づけることが可能であるかもしれない、もしこの関連が発見されれば情報学の急速な発展が期待できること、数学それ自体は余りに一般的であり、オペレーショナルリサーチそれ自体は余りに個別的事であること、等々、関連諸領域を具体的にあげてそれらの貢献を否定しながら、全体としては、ミハイロフらに同意して、“それゆえ、過去5年間にわたる理論インフォーマティクス⁴⁾の発展に関するレポートを依頼された人は誰でも編集者に『皆無』という報告、すなわち『未だ理論はありません。それゆえ何の進歩もありえません。』という報告、を送れるし、それで不当ではない、”⁵⁾と述べている。しかし、情報学の理論的基盤を強固にする可能性を持つ唯一の法則として、科学者の生産性、雑誌と関連論文数、単語の使用頻度について、それぞれロトカ、ブラッドフォード、ジップが得た経験則を説明し、この項を次の言葉で結んでいる：“その中にひそむメカニズムがいつそう完全に理解されたとき、インフォーマティクスのこれらの諸法則は他の社会諸科学にも応用されるであろう。”⁶⁾ 彼は、他の論文の中でも、“書誌の数量的分析の主要な目的”の一つとして、“情報学の理論を発展させる基礎を形成しうるのである経験則の発見と解明”⁷⁾を挙げている。

2. システム分析

企業体におけるシステムの思考法の重視は、生産規模の拡大、生産性の向上、営業範囲の拡大、他システムとの相互作用の増大、これらによる情報の増加、そして社会変化の急激性など、企業の存続・成長に影響を与える

さまざまな要因がますます複雑になり、以前のような勘による経営ではたちゆかなくなったことによる。そして、図書館・情報システムの管理運営にもシステム分析⁸⁾の手法が適用されねばならないとする主張の根拠も、記録された知識の量の増加、その質の多様化、情報流通の組織化の必要性の認識、利用者の要求の多様化などによる図書館・情報システムの増加、大規模化、複雑化がある程度進んだ段階で予算の制約がますます強まり、図書館も、利潤の増大をはかる必要はないとしても、少なくとも業務の効率という点においては、営利企業と同じ立場に立って、コスト/イフェクティブネス、コスト/パフォーマンスなどを考えざるを得なくなったこと、に求められる。こうした考え方は、図書館業務を処理するためにコンピュータが導入されるようになったとき、伝統的な図書館業務が再検討されることによって広まった。

しかし、システム分析は、他で得られた結果をそのまま適用できるものではないという点で、他の自然科学の成果とは異なる。人間の作りあげた組織は環境の影響を最も受け易いものであるし、まったく同じ目標を持つ組織はないからである。システム分析に関する研究が与えることのできるのは問題解決のためのツールあるいは方法論だけであり、あらゆる要因の抽出と全体の調整が最も重要な仕事となる。⁹⁾ システム科学とは、システム効率化のための方法論の定式化・図式化であるということもできよう。

システム分析を行なう場合は、「システムの目標の設定」がまず行なわれなければならない。この目標はその後若干の修正を受けることもあるが、これなくしてはシステムの設計も評価もできないことになる。企業の目標は、一般に、利潤の最大化とすることができる。しかし、図書館・情報システムは、図書館情報学の基礎理論が図書館・情報システムの本質的機能の理解に到達したといえるほど充分に発展させられたとはいえないので、明確な目標を設定することが困難である。「利用者の満足度を最大にするサービスの提供」といった抽象的な目標にはその満足度をはかる具体的な尺度が必要とされようし、「単位時間あたりの処理可能利用者数の最大化」といった目標は図書館の目標とはほど遠いものである。しかし、このことは必ずしも、図書館情報学の基礎理論が確立されるまでは図書館・情報システムは目標を設定しえない、ということの意味しない。例えば、公共図書館においては、「貸出冊数・登録者数を可能な限り最大にする」という目標は、第一歩としては、一つの目標とな

りうる。

システム分析は方法論であるから、関連データは対象とするシステムのそれぞれについて収集しなければならない。そして、一見同じように見える研究でも、それぞれ独自の意義を主張しうる。このことが多くの論文が発表される理由である。図書館に関わる諸局面の計数作業は、これまで、図書館統計・利用者統計と呼ばれて、ある程度までは行なわれていた。しかし、この統計は、そこから図書館経営の指針を得ることができるほど充分なものではなかった。すなわち、データの収集は行なわれても、多くは習慣として行なわれている目的なき収集であるため、それらのデータが目的を持った分析の有用なベースとなり得ることはきわめて少ない、いわば統計のための統計であった、といえる。¹⁰⁾ そして、よりミクロな面に関するデータは我々はほとんど持っていない。モースは、図書館のコレクションの利用に関する統計を例に引いて、将来の利用を予測するためには過去の利用のデータが不可欠であるのに図書館はこのようなデータを持っていないことを批判している：「図書館員がコレクションの各クラスの図書の各年度の平均貸出率に関する数字を持っていないということ、そして、わずかの最も利用度の高い図書についての個々の貸出率に関する数字を持っている場合はもっと少ないということ、に対して図書館員は弁解の余地はない。」¹¹⁾

ところで、データ収集は個々のシステムに対して行なう必要があるということは、これら個々のデータの組、あるいはそこから得られた平均値、分散、メジアン、モード、範囲、等々の統計量の比較しかできないことを意味するものではない。システム分析では、主としてオペレーションズリサーチの分野で開発された手法を応用して、数学モデルがつくられることが多い。そして、ひとたび、ある現象のふるまいがある数学モデルで記述できることが確認されると、それと類似の現象のふるまいは、パラメータを決めるに必要なわずかのデータを収集するだけで、予測できるようになる。

ビブリオメトリクスの現段階は、データの収集と並べかえ、グラフ化、経験式の導出までで、この式のパラメータは個々の書誌の諸変量といかなる関係にあるか、ある分野の文献の特性を表わすパラメータは何かはわかっていない。いくつかのパラメータが提案され、一応の意味づけがなされているが、異なる分野間の文献の特性の比較に耐えるほど説得力のあるパラメータはまだない。このため、まず数学モデルの正しさを確認しようとして、

すなわち、ブラッドフォードの法則はいずれの分野でも成り立つ法則であることを確認しようとして、また各分野の重要雑誌を識別しようとして、多くの分野について研究が行なわれた。

システム分析の一方法として、ビブリオメトリクスの手法を応用した研究の例として、OSTI の補助金を受けてランカスター大学のグループによって行なわれた、大学図書館のシステム分析の研究がある。¹²⁾ 効率的な図書館・情報システムの設計における、雑誌の利用度・寿命などに関する研究の有用性については後にふれる。

3. 利用者研究

現在、図書館情報学の分野で、多額の研究費が投じられ大規模な研究が行なわれているのは、機械化と利用者に関する研究であるといつてよい。機械化の研究が多いのは、計算機の使用を必要とする情報サービスの増加と、事務処理のための計算機導入の流行とによる。また、利用者研究が盛んになった理由は、図書館が、急激に変化する社会の中で、みずからの提供するサービスと利用者の要求との間に懸隔が生じてきたことを自覚し（あるいは、それに対して危機の念を抱き）始めたこと、多くの地域的・全国的システム建設の計画が生まれ、これらを利用者のニーズにより適合するシステムにするためには利用者の特性を把握する必要があったこと、またこれに対して政府機関から多額の調査研究費が与えられたこと、いくつかの学会などが学術情報流通システムの抜本的な改善の必要性についての調査を始めたこと、などに求められる。¹³⁾

利用者研究は主として面接法、質問紙法、日記などによって行なわれるが、科学文献の分析も有力な方法として利用されている。しかし、これは補助的な方法でしかないようである：「体系的アプローチ（コミュニケーションの物的媒体による研究）は時間的に変化していくコミュニケーションのパターンの研究に特に適している。そして、ここでこれをとりあげるのも、この意味においてである。体系的アプローチは情報利用のパターンおよび情報要求に関するデータを与えるが、情報に対するニーズに関するデータを提供することはできない。充足されなかった情報ニーズはコミュニケーションの物的媒体とは結びついておらず、それゆえ体系的アプローチによっては識別することはできない。」¹⁴⁾

利用者研究もニーズを知るための間接的手段でしかなく、また現在行なわれている利用者研究に対する批判も少なくないが、¹⁵⁾ このような批判は本論の目的ではな

い。ここでは、利用者研究の調査結果はビブリオメトリクスで利用できるデータを多く含んでいるということ、よりよい図書館・情報システムの建設のためにはビブリオメトリクスの手法が充分活用されるべきであること、にだけ注意しておこう。利用者研究の論文も 1965 年頃から急激に増加した。

III. 歴 史 的 回 顧

本章では 1960 年代の中頃までに行なわれた研究を扱う。これらは個別的な研究で、そこから一般的な手順・法則が導かれたことはなく、方法も画一的であった。唯一の例外として、ブラッドフォードが 1934 年に発表した法則があるが、その重要性が広く認められるようになったのはごく最近である。それゆえ、この法則は後の章で扱いたい。

1. 統計的書誌学

1) コールとイールズの研究

ビブリオメトリクスに関する研究で最も古いと思われるのは、コールらによる比較解剖学の分野における文献数の歴史的变化の研究である。彼らは 1543 年から 1860 年までに発表された動物の解剖に関する論文について次のような分析を行なった。

- (1) 16 世紀から 1860 年までの比較解剖学者全体の活動を曲線によって表わす。
- (2) この全般的な変化を分解して、ヨーロッパ各国の業績を別々にプロットする。
- (3) 同様な方法で、各時代に、どのグループの動物およびその主題のどのような局面が研究者の注意をひいたかを知る。
- (4) 各時代の事件、公的団体および個人が解剖学思想の歴史に与えた影響を跡づける。¹⁶⁾

このような方法は、論文の数しか考慮せず、その質は一切問わないという欠点を持つ。しかしそれにもかかわらず、この方法は次のような意義をもっていると彼らは考えていた：“図から得られたいかなる結論も、扱われている論文の科学上の価値を検討することによって、チェックしなければならない。しかしそれでもなお数値のみでも価値を有するのだということが出来る。ある特定の時代に多くの研究が発表されたという事実は、その大部分がいかに平凡なものであっても、その時代の関心と活動の確かな指標となっている。”¹⁷⁾

数量化によるある程度の質の無視は（それを補正する方法はいくつかあるが）計量的方法の避けがたい限界で

あるが、彼らの研究で評価されるべき点は、国別、主題別（魚類及び下等脊索動物、両棲類、爬虫類、鳥類など）の論文数の年次推移を、大科学者の活躍、学会の設立、研究成果の発表方法の変化、研究テーマに対する関心の移り変わりなど、科学研究を規定するさまざまな要因との関連で考察したことにある。

2) ヒュームの統計的書誌学

コールらの「文献の統計的分析」は比較解剖学の歴史を書くための基礎作業として行なわれたものであった。これに対して、科学文献全体の数の変化から、科学・社会・産業の発展の歴史を跡づけようと試みたのはヒュームである。¹⁸⁾

彼は、書誌のデータが現代文明の進歩における諸変化を例証し解釈するための補助として重要であることを説き、*International Catalogue of Scientific Literature (ICSL)* および *Catalogue of Scientific Papers (CSP)* を利用していくつかの調査を行ない、その結果を図や表にまとめて、それまで他のソースにより得られていた結論を補強したばかりでなく、社会・産業の歴史における重要な動きをより正確に説明することに成功した。¹⁹⁾

彼の研究は、*ICSL* の対象が純粋科学に限られていたため技術の進歩の与えた影響が表面には現われない、この目録が採用した 17 分野の分類にそのまま従い、各分野を一つの単位として計数したため、より狭いテーマに関する研究の発展・衰退またはテーマ間の相互関係などはわからない、などの欠点を有していたにも拘わらず、多くの問題を解明するために、書誌学者や統計学者が発展させるべき価値のある内容を持つものとして評価された。

ヒュームは自分の行なった研究を「統計的書誌学」(Statistical Bibliography) と名づけ、こうした研究の重要性を強調したが、彼自身は研究をつづけず、あとにつづく者もいなかった。最大の原因は、彼がこの講演を行なった年に、*ICSL* の刊行廃止が決定され、最も便利なツールが利用できなくなったことにあると思われる。これ以後、*Science Since Babylon* にはじまる一連の著作によって書誌の数量的分析の有効性がブライスにより実証されるまでは、このような研究はほとんど行なわれなかった。²⁰⁾

3) ロトカの法則

発表論文数の時代的推移ではなく、科学文献とそれを書く科学者との関係に注目し、科学者の論文の生産性に関する一般的法則を導びいたのは、形式人口学における

安定人口論で有名な、ロトカである。²¹⁾

彼は、*Chemical Abstracts* の累積索引 (1907-1916) および *Auerbach's Geschichtstafeln der Physik* の著者索引によって、すべての著者 (ただし、*Chemical Abstracts* についてはイニシャルが A, B の科学者のみ) の発表論文数を数え上げ、発表論文数とそれぞれの数の論文を発表した科学者の数との関係を両対数グラフ上にプロットした。このグラフ上で両者の関係はほぼ直線で近似できること、すなわち、 x 篇の論文を発表した科学者の全著者数に対する割合を y とすると、 x と y との間には、

$$x^n y = \text{const.} \quad (1)$$

なる関係が成立すること、また最小二乗法により求めた n の値はいずれの場合も約 2 となることなどを彼は見出し、これを「科学の生産性に関する逆二乗則」(Inverse Square Law of Scientific Productivity) と呼んだ。今日、これは「ロトカの法則」(Lotka's Law) として知られている。

この法則について彼は「一般形で表わした度数分布 (1) はさまざまな現象に対して広い応用範囲を持っている。そして単にそのような形の分布をしていると言うだけでは、その中にひそむ物理的関係にはほとんどあるいはまったく光をあてない。」²²⁾ と述べて、この法則は科学の制度の中にある何らかの「物理的関係」によって生まれており、この関係の解明にまで進むべきであると考えていたようである。

彼の言う「物理的関係」とは、その時代の科学研究システムあるいは論文発表システムに加えられているある種の強制力によって作られている関係と考えることができる。自然科学、人文科学²³⁾ を問わず、いつの時代にも、学問研究の世界において活発な活動をする科学者がおり、このような人々が科学研究において重要な役割を果たしている事実は、科学の社会システムを解明する重要な手がかりとなり、それゆえ、「科学の社会学」の恰好の研究テーマとなっている。

2. 引用文献調査

このような研究に少しおくれで、図書館学の分野では、主として図書館における雑誌選択という実的な問題を解決するための方法として、一次雑誌 (primary journal) の引用文献の掲載誌を調べることによって雑誌の「重要度」を知ろうとする試みがなされた。この研究は数多くの分野について行なわれ、研究例はかなりの数にのぼる。これらの個々の研究についてはいくつかの論文の中

で紹介されているので、²⁴⁾ この節の記述は、個々の研究の紹介よりは、むしろ、どのような問題を解明するために引用文献調査が行なわれたかに重点をおきたい。

1) タイトルの分散

タイトルの分散 (title dispersion) とは、ある特定の主題に関する論文が、その主題を中心として扱っている雑誌ばかりでなく、一見関連のなさそうな他の分野を扱っている雑誌にも多く掲載されていることをいう。この分散の様子は、現在、「ブラッドフォードの法則」として知られており、ビブリオメトリクスの主要なテーマの一つとなっている。²⁵⁾

図書館における図書選択および蔵書評価の重要性は古くから認識されており、そのために 1920 年代の中頃までには図書選択の主要な理論はほぼ出そろっていたが、雑誌については、図書と同様の方法によって選択する以外、雑誌の特殊な性格を考慮に入れた固有の選択方法はなかった。雑誌論文の比重が図書ほど高くない分野においては重要な雑誌の識別はそれほど困難ではないであろうが、すでに研究成果の伝達手段として雑誌が著しく重要な地位を占めていた自然科学の分野で雑誌の選択基準が求められるのは当然であろう。そして、さらに、自然科学の中でも進んだ分野の一つであった化学において、雑誌選択の具体的方法が考え出されたのは偶然ではない。タイトルの分散に関する研究は、このような雑誌の重要性の高まりと学部教育から大学院教育への重点の移行という当時の大学教育の変化を背景として、この変化の重要性を感じとった大学教師グロスらによって、学生たちをこのような変化に対応できるように効果的に教育するためのステップとして「教授陣に対する刺激および彼らの知的発達に必要な資料をも考慮に入れつつ、学生が大学院における研究にうまく進んでいくための準備として、カレッジ図書館においてはいかなる科学の定期刊行物のコレクションが必要とされているか」²⁶⁾ を知るために行なわれた。

もちろん雑誌についても、図書の場合と同じく、リストを作成し、これに基づいて選択することもできるが、この場合、編纂者の専門領域・好みによる偏りは、図書のリストよりは、はるかに大きなものになってしまうだろう。これより、教師が、学生たちに化学を理解させ、その進歩に貢献できるように学生たちを訓練するためには、現在その分野で研究している人々がどのようなツールを使っているかを調べるべきである、とグロスらは考えた。この考えに基づいて彼らはアメリカ化学会の機関

誌 *The Journal of the American Chemical Society* の 1 巻 (1926 年刊行分) に掲載された論文に引用されているすべての参考文献の掲載誌を調べ上げ、それぞれの雑誌が引用された回数を発行年度別に表にまとめた (ただし、*J. Am. Chem. Soc.* は除かれていることに注意)。彼らはこれより、総引用回数および最近 10 年間の引用回数の多い順に雑誌のタイトルを並べたリストを作成し、この二つのリストに基づいて、化学関係の図書館で備えるべき雑誌のタイトルと必要なバックナンバーを挙げた。

この後につづく、引用文献調査による、タイトルの分散に関する研究はすべて、この著者たちの「化学以外の分野の研究者に対して彼らが彼ら自身のためにこの (いかなる雑誌がカレッジ図書館において必要かという) 問題に解答を与える方法は示した」⁽²⁷⁾ という言葉を受けて、彼らとほとんど同じ方法を他の分野に適用したにすぎない。この間に改良された点は調査対象雑誌を 1 タイトルではなく 2 タイトル以上とする方法が考え出されたことだけであり、この意味では、引用文献調査の方法はグロスらによってほぼ確立されたと考えられることもできよう。しかし最近ではさまざまな方法によってタイトルの分散が調べられている。これについては後にふれる。

彼らの方法はたしかに科学的ではあったが、雑誌選択における採否の境界が明確に定められていなかったことが一つの難点としてあげられよう。彼らは「いずれの場合も 1 年あたりのコストは同程度と仮定すると、創刊後 10 年しか経過しておらず、かつ 50 年間続いている雑誌の、例えば、半数の引用文献しか含まない雑誌でも、投下する資金から見ると、この雑誌の方が、50 年間続いている雑誌よりも望ましいということが認識されなければならない」⁽²⁸⁾ と述べて、実際の選択にあたっては、コストも考慮に入れなければならないことを強調しているが、彼らが必要であるとして挙げた雑誌はコストとの関連で下限が定められたわけではなく、またそれ以外の何らかの基準によったものであるともいっていない。その下限は、やはり、化学者の直感によって定められたとしか考えられない。

2) 時間的広がり

ある論文が発表されてからこの論文を引用した論文が発表されるまでに経過した年数 (ふつう小数点以下第一位までとれば充分である) をその論文の年令と定義する。時間的広がり (temporal span) とは引用文献の年令分布を言う。現在これは「文献の寿命」、文献の「半

減期」(half-life)、あるいは文献の「利用度減少」(obsolescence) などと呼ばれている。

「年をとった」論文ほど利用度が少なくなるという事実はすでにグロスらの引用文献調査によって明らかにされていた。彼らが購入すべきバックナンバーを決めるという点から利用度減少を考えたのに対して、研究図書館の建物の効率的な利用、すなわち資料の廃棄・退蔵という今日的な観点からその重要性を指摘したのはフスラーである。⁽²⁹⁾ 彼は「研究図書館は余りにも急速に成長しつつあるので効率的な収容ができなくなっている」⁽³⁰⁾ ことを重視し、利用度減少に関する研究が何らかのコントロールを加える際の尺度となりうると考えた。すなわち、それまでの図書館はどの資料でもいつかは誰かが利用するに違いないという仮定に立っていたため、廃棄 (discarding) あるいは別置 (segregation) は実際には困難であったが、将来の研究図書館は、いつかは利用されるという可能性よりは、「注意深く評価を加えられた確率に基づいた運営がますます必要になってくるだろう」⁽³¹⁾ と述べて、科学的方法の重要性を強調した。

利用度減少は、とくに図書館の管理運営を考える場合に、文献の成長と並んで、ブラッドフォードの法則と密接な関連があるので、これについては IV でさらに詳しく扱う。

3) 言語の分布

言語の分布 (language distribution) とは、引用文献の総数に対する各国語で書かれた文献の割合をいう。

科学論文がラテン語で書かれ、また大学における講義もラテン語で行なわれていた近世初期までは、ラテン語が唯一の科学の言語であった。この慣行が初めて破れるのは、16 世紀に、ロンドンのグresham・カレッジで英語によって講義が行なわれた時である。しかし、研究成果の主たる発表手段であった科学書は依然としてラテン語で書かれており、多くの科学論文・科学書が自国語で書かれるようになるのは 18 世紀以降のことである。この慣行が定着すると、科学者は、いくつかの外国語を、その重要度に応じて、学ばなければならない。そして言語の分布に関する調査も重要な意義を持つてくる。

言語の分布の調査もグロスらによって初めて行なわれた。⁽³²⁾ 彼らはその結果 (ドイツ語 52.5%, 英語 35.2%, フランス語 9.4%) から、カレッジで化学を専攻するすべての学生はドイツ語の読解力をつける必要がある、フランス語がそれに代わりうるとはほとんど認められない、という結論を得ている。しかし、1920 年代あるい

は 30 年代における言語の分布は、現在のそれとはかなりの違いがある。現在、引用文献に英語の論文が圧倒的に多いのは、米国が学問研究の中心となっていること、英語が国際語としての地位を確保したこと、またそれと相互的な関係にあるが、英語国以外の多くの国が学校教育の中で英語を第一外国語として定めていることなどによっていると思われるが、最大の理由は、これらの相乗効果による英語の重要度のいっそうの増大と英語以外の言語の軽視によるものと考えられる。このため、言語の分布に関する調査結果は、今日、少なくとも引用文献による限りでは、大学において修得すべき外国語を決める際の判断基準として利用するには不十分な点が多く、言語の重要度の時代的推移により科学の繁栄の中心地を知るといった歴史的興味以上のものを与えてくれそうにない。言語の分布の時代的推移と各国の研究投資額・科学者数の推移、産業の発展などの相関からは興味ある結論を引き出すことができよう。³³⁾

4) 主題の分散

主題の分散 (subject dispersion) とは、ある分野の論文は各分野を扱っている雑誌に発表された論文をどの程度引用しているかをいう。この調査は、ヴォイトにより、工学あるいは農学の研究には純粋科学の雑誌がどの程度必要かを知るために行なわれた。³⁴⁾ しかし、ある雑誌に発表された論文の主題は必ずしもその雑誌のタイトルから想像される分野に属するとは言えないわけであるから、原論文および引用論文の属する分野を特定した上で、ある主題の文献がどのような分野の文献を引用しているかを調べるならこれは学問領域間の相関について有用なデータを提供すると考えられるが、彼の行なったのは「どの雑誌に掲載された論文を多く引用しているか」の調査なので、とくに具体的に図書館業務に応用するときには個々のタイトル名が必要とされることを考えると、雑誌の配架の参考にもするならともかく、タイトルの分散とはなれてこの調査を行なう意義ははっきりしなくなる。

このように雑誌のタイトルを一つの単位として考えることにはある程度の限界があるが、その限界を考慮に入れても、いくつかの興味ある事実を知ることができる。例えば、フスラーは、主題の分散を文献の利用度減少と結びつけ、引用文献が発表された雑誌の分野の違いによる文献の利用の遅れを調べた。³⁵⁾ 彼は、“化学者は、物理学の雑誌に発表された論文よりは、化学の雑誌に発表された論文をはやく利用するであろう”³⁶⁾ という仮説を

設定し、それぞれの分野の雑誌の引用文献の年令のメジアンとして次のような結果を得て、この仮説の正しいことを示した。

表 1. 物理学者と化学者との間に見られる雑誌利用の遅れの違い (1939 年)³⁷⁾

	化学関係雑誌	物理学関係雑誌
化学者	5 年	7.5 年
物理学者	5	3

このような遅れの違いは、もちろん研究テーマの特殊性にもよるが、主として、一人または少数の研究者による「他分野」の雑誌中の論文の発見と紹介、その後の利用の増加という他分野の文献の利用における媒介機構の存在によって生じていると考えられる。この差は情報サービスおよび図書館の充実によってかなり小さくなると思われるが、いかに充実した情報サービスが行なわれても、他分野の研究成果を自分の研究に応用して成果をあげることは少数の優れた科学者にしかできないことであるから、この差はなくならないと考えられる。そして、それゆえに、科学者の特性を知るためには興味のあるテーマである。³⁸⁾

5) 文献の形態

文献の形態 (form) とは、引用文献が、雑誌論文、モノグラフ、図書、学位論文、あるいは特許明細書などのうち、いかなる形で発表されたかをいう。文献の形態に関する調査もグロスらによって初めて行なわれたが、³⁹⁾ 今日でもなお、図書選択および雑誌選択についてはそれぞれの方法がある程度確立されているにも拘わらず、図書と雑誌の相対的な重要度の指標を得る方法は何もない。このため、図書館における資料購入費の配分なども、直感によるか、他人にまかせるしかなく、たとえ図書館によってこの配分の比率が決められたとしても、それは一貫した理論に基づく方法によっているとは考えられない。もちろんこの配分を決める際には、それぞれの形態をもつ文献の利用度数、利用の性格の相違などを考慮に入れなければならないが、文献の形態に関する調査はこの配分を決める際の手がかりとして決して軽視されてはならない。とくに、同じ自然科学でも分野によっては雑誌が最も重要な情報源ではない場合もあるので、⁴⁰⁾ この調査も、タイトルの分散に関する調査と同じくらいの頻度で、くり返し行なう必要がある。

しかしながら、この調査の意義は、図書館業務に有効

な方法を提供することにとどまらない。科学におけるコミュニケーションの手段として300年以上の歴史をもつ科学雑誌は、現在、論文の出版の遅れ、論文の長さの制限、関連文献の多くの雑誌への分散、雑誌の価格の高さ、科学者が編集に費やす時間の無駄などを中心とする多くの批判にさらされている。この結果、一方では、第一線の科学者を中心とする現代の「インヴィジブル・カレッジ」が形成され、他方では、科学雑誌に代わるべき科学のコミュニケーションの手段が、アメリカ心理学会などによって、模索されている。このような事情を背景として、現在、モノグラフの重要性がいっそう増大しており、⁴¹⁾ これはとくに変化の激しい分野（このような分野はタイトルの分散度も大きい⁴²⁾）において著しい。研究成果の発表がどのように行なわれているかを知ることが、科学雑誌の将来を考える上で、無視できないことからである。

IV. ブラッドフォードの法則

1. ブラッドフォードによる法則の定式化

1) 実験の背景⁴³⁾

前章で述べたタイトルの分散に関する研究は、現在、引用文献調査によると否とを問わず、「ブラッドフォードの分散則」(Bradford's Law of Scattering)あるいは「ブラッドフォードの法則」(Bradford's Law)と呼ばれて、ビブリオメトリクスの中心的なテーマとなっている。そして、タイトルの分散に関する研究のほとんどすべてが図書館における雑誌選択のために行なわれたのと同様に、ブラッドフォードの法則と呼ばれている今日でも、その研究のほとんどは雑誌選択の指針を得る目的で行なわれている。しかし、ブラッドフォードがこの法則を定式化するもとなる実験を行なった動機は、これとは別のところにあった。

科学者としての教育を受け、実際に科学研究に従事したこともあるブラッドフォードは、科学の進歩は記録された知識に依存し、この知識に導びくのは抄録・索引誌などの二次資料であり、それゆえ、研究者の間によく見られる研究の重複は、彼らの研究している主題に関連のある重要な論文を、抄録・索引誌がとりあげていないことがしばしばあることによる、と考えていた。彼は、このような書誌サービスの不完全さに起因する、研究に必要な科学文献への接近の困難さを指して、「文献の混沌」(documentary chaos)あるいは「書誌の混沌」(bibliographical chaos)と呼んだ。⁴⁴⁾

彼は、主として彼が当時館長であった英国科学博物館図書館 (Science Museum Library, SML)⁴⁵⁾ で受け入れている雑誌を利用して、当時の書誌サービスがどの程度不完全であるかを調べた結果、「発表された有用な科学論文のうち、抄録誌に収録されるものは半数にも満たず、残りの半数以上の有用な発明・発見は、発表されても何の役目も果たさず注意もされないで、図書館の書架に置かれているだけである」⁴⁶⁾ との結論を得た。これはサンプル調査に基づく推定であったので、これを確認するために、電気工学の分野を例にとりて、詳細な調査を行なった。その結果、電気工学に関連のある21の抄録・索引誌によって到達しうる電気工学の論文は半数以下であり、さらに、最も多く論文を収録している索引誌でもその数は全論文の5分の1でしかないので、有用な論文の半数以下を見つけるためにさえ21の抄録・索引誌のすべてにあたらなければならないことが明らかになった。これは、一つには、「個々の抄録機関は自分の受け持っている分野に関連のある雑誌およびこれらと密接な関連のあるわずかの分野の雑誌に掲載された論文にしか注意を払わず、「他の」雑誌に掲載された関連論文の抄録はその雑誌の属する分野を受け持つ抄録機関にまかせてしまっている」ためであり、もう一つには、「抄録機関は、一般に、図書、パンフレット、特許明細書、等々の資料を無視してしまっている」⁴⁷⁾ ためであった。

これら二つの原因のうち、第二の原因は抄録誌の収録範囲を他の形態の文献にまで拡大することによって取り除くことができるが、第一の原因の詳細、すなわち関連文献の分散の程度については何も知られていなかった。これより、完全なドキュメンテーションを行なうには、ある主題に関する論文が他の主題を扱っている雑誌にどの程度あらわれているかを調べるが必要になる。

ところで、ある分野の雑誌には、それ以外の分野を扱っている雑誌よりも、その分野の論文が多く掲載されていることは明らかである。このため、雑誌を、ある主題に関する論文を多く載せているという意味で、生産性の高い順に並べることが可能であり、この時、同じ数の論文を含むように雑誌をいくつかのゾーンに分けると、各ゾーンの雑誌の数は生産性が減るにつれて増すであろう。これが彼の設定した仮説である。

2) 法則の定式化

この仮説を検証するために彼がとった方法は、それ以前から行なわれていた、引用文献による方法ではなかった。⁴⁸⁾ 彼は、「応用地球物理学」(applied geophysics)

と「潤滑」(lubrication)の二つの主題を選んで、これらのカレントな書誌を作成し、⁴⁹⁾ この二つの主題書誌に採録された論文とその掲載誌を調べて、例えば応用地球物理学について、表2のような結果を得た。

表2. 主題書誌「応用地球物理学」に収録された論文と掲載誌との数量的関係 (1928-1931 年)⁵⁰⁾

A	B	C	D	E
1	93	1	93	0.000
1	86	2	179	0.301
1	56	3	235	0.477
1	48	4	283	0.602
1	46	5	329	0.699
1	35	6	364	0.778
1	28	7	392	0.845
1	20	8	412	0.903
1	17	9	429	0.954
4	16	13	493	1.114
1	15	14	508	1.146
5	14	19	578	1.279
1	12	20	590	1.301
2	11	22	612	1.342
5	10	27	662	1.431
3	9	30	689	1.477
8	8	38	753	1.580
7	7	45	802	1.653
11	6	56	868	1.748
12	5	68	928	1.833
17	4	85	996	1.929
23	3	108	1,065	2.033
49	2	157	1,163	2.196
169	1	326	1,332	2.513

A: B 欄に示す数だけ論文を掲載している雑誌の数
 B: 関連論文数
 C: A 欄の数の上からの累積和
 D: A と B の積の上からの累積和
 E: C 欄の数の常用対数.

この、E, D の関係をグラフに描くと、図1のように、応用地球物理学、潤滑のいずれについても、傾きは異なるが、同じような形の曲線が得られる。⁵¹⁾ 図2はこれを一般化したもので、この図で r に含まれる雑誌がその分野の核となる雑誌である。また α , β , γ の間には、

$$\alpha: \beta: \gamma = 1: n: n^2 \quad (2)$$

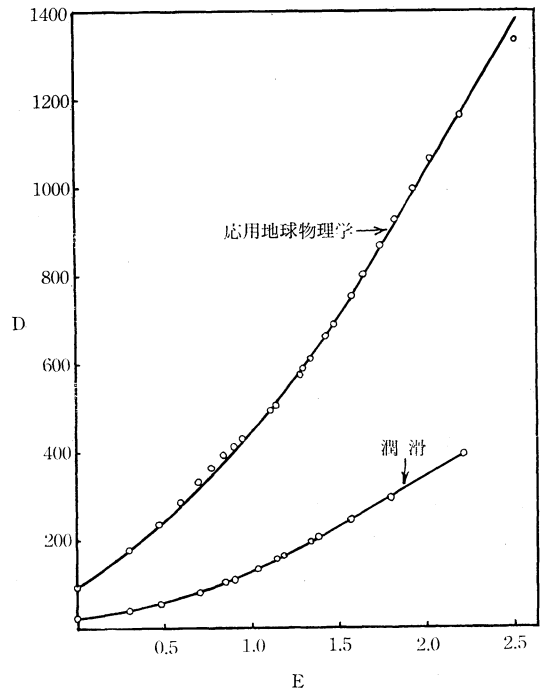


図1. 「応用地球物理学」および「潤滑」に関する論文とその掲載誌との数量的関係

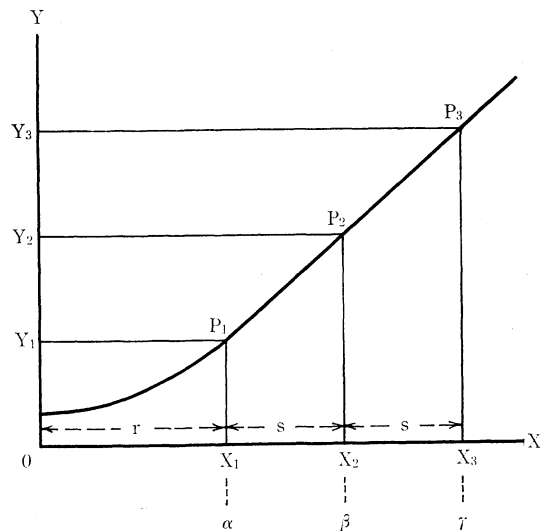


図2. 特定の主題に関する論文とその掲載誌との一般化された数量的関係

なる関係がある。これがグラフによって表わされた「ブラッドフォードの分散則」⁵²⁾ である。

ところが、彼は、この図に基づきながら、科学論文の分散に関する法則を次のように述べた。

科学雑誌をある与えられた主題に関する論文の生産性の高い順に並べた場合、それらの雑誌は、他の雑誌よりはその主題を中心に扱っている核 (nucleus) と、その核と同数の論文を含むいくつかのグループあるいはゾーンとに分けられる。この時、核およびそれに続くゾーンに含まれる雑誌の数 (の比) は $1 : n : n^2 : \dots$ となる。⁵³⁾

これが文章によって表わされた「ブラッドフォードの分散則」である。これより、彼は、「数多くの論文が、アブリオリには「載っていきそうもない」雑誌に掲載されている⁵⁴⁾ という一般的結論を導びき出した。そして彼は、論文の分散におけるこの規則性は理論的には「科学の統一性」 (unity of science) から演繹され、「この原理によれば、あらゆる科学の分科は、その親近性の程度に違いはあるが、他のすべての分科と関連している」⁵⁵⁾ と説明している。なお、nucleus は現在 core journals, key journals, key-periodicals, 'star' periodicals などとも呼ばれている。

ところで、ここで述べた法則の二つの表現、すなわち「グラフによる表現」と「文章による表現」、は実は同じ内容ではない。これについては次節でふれることにして、その前に、「分散」 (scattering) という言葉について若干説明を加えておきたい。ガーフィールドは、「ブラッドフォードの分散則」の一般化であるとして、「ガーフィールドの集中則」 (Garfield's Law of Concentration) なる法則を提案している。⁵⁶⁾ この「法則」は、関連論文が多くの雑誌に掲載されていることを、逆に、その大部分はわずかの雑誌に集中していると述べたものであるように受けとられやすい。⁵⁷⁾ しかし、彼の「集中則」は、「分散則」を視点を変えてながめて、提案された法則ではなく、「雑誌は、基本的には、すべての分野に共通のコアあるいは核に集中している」⁵⁸⁾ ことをいっている。すなわち、お互いに無関係ないくつかの分野を代表する図書館員たちが、それぞれの分野について調査した結果つくりあげた、500 から 1000 誌の雑誌リストを互いに交換し、他の分野のリストに基づいて雑誌を購入してコレクションをつくりあげても何もおこらず、「事実、各々の図書館員は「彼の持っている」リストが彼の専門図書館のニーズにみごとに適合するコレクションをつくりあげたことを見出す」⁵⁹⁾ ことを意味してい

る。しかし、総合科学図書館をつくるのなら話は別であるが、また現在世界中で発行されている雑誌は三万タイトルとも五万タイトルとも言われているがそのうちでよく利用される雑誌はごく一部でしかないということをおもうとするなら話は別であるが、限られた予算でできるだけ多くの関連文献を利用できるようにしたいと願う図書館員にとって、1000の雑誌を集めればどのような分野の専門図書館のコレクションとしても充分であるなどという「法則」は何の役に立つであろうか。もし、これを上位50誌に限れば、それぞれの分野についてつくられたリストの間には、分野の関連性がうすくなると、ほとんどのリストに現われるであろう *Science*, *Nature* など一部の総合科学雑誌を除くと、共通のタイトルが見られなくなることは明らかである。⁶⁰⁾

このように「分散則」と「集中則」とはその内容を異にするものであるが、もし仮に、この二つを同じ現象を別の視点からながめたものと考えても、「ブラッドフォードの分散則」が「集中則」と名づけられることはなかったと思われる。たしかに、図書館の蔵書構成を考える場合、あるいは二次資料を作成する場合でさえ、いっその注意を払うべきは主要雑誌の識別とその収集であり、その限りにおいては、文献の分散という事実を文献の「集中」として扱っても何ら不都合は生じない。しかし、あらゆる関連論文の一つものがさない二次資料の作成、すなわち「完全なドキュメンテーション」をみずからの図書館・ドキュメンテーション活動の使命と考えていたブラッドフォードにとって、その関心が、当時の抄録・索引誌でも十分にカバーしていた核雑誌ではなく、一般には載っていないと考えられている雑誌にも関連論文が載っているという事実の方に向かうのは当然であった。彼が「分散」ということばを使ったのもこのような事実に対して注意を喚起するためであったと考えられる。

2. 法則の数学的表現

1) 法則の数学的表現

まず、いくつかの記号を定義しておく。ここで、雑誌は論文の生産性の高い順に並べられているものとする。

n : 雑誌の順位 (同じ順位の雑誌が2以上ある場合は、1をこえる分だけ次の順位を空ける)

$r(n)$: n 番目の雑誌に含まれる関連論文の数

$R(n)$: 上位 n 誌に含まれる関連論文の数、すなわち、 $r(n)$ の $r(1)$ から $r(n)$ までの累積和

N_r : 上位 r 篇の関連論文を含む雑誌の数

N : 関連論文を1以上含む雑誌の総数

さて、「グラフによる表現」と「文章による表現」とは矛盾した内容を持つものであることを初めて指摘したのはヴィッカリーである。⁶¹⁾ 彼は、「グラフ」の場合は、雑誌を上位からとったとき、ある数の論文が得られる雑誌の順位の間に成り立つ関係であるのに対して、「文章」の場合は各ゾーンに含まれる雑誌の数の間に成り立つ関係であることを指摘し、これら二つの表現はそれぞれ次のように表わされることを示した。ただし、 c , b は定数である。

グラフによる表現

$$N_r : N_{2r} : N_{3r} : \dots = 1 : a : a^2 : \dots \quad (3)$$

文章による表現

$$N_r : (N_{2r} - N_r) : (N_{3r} - N_{2r}) : \dots \\ = 1 : b : b^2 : \dots \quad (4)$$

ところで、ブラッドフォードの法則は、一般に、雑誌の順位とその順位までの雑誌から得られる関連論文数との間の関係として表わされる。このため、上式を n と $R(n)$ との関係として表わすと、

グラフによる表現⁶²⁾

$$R(n) = k \ln \frac{n}{s} \quad (5)$$

文章による表現⁶³⁾

$$R(n) = a \ln \left(\frac{n}{b} + 1 \right) \quad (6)$$

となる。⁶⁴⁾ ただし、 k , s , a , b は文献によって変わるパラメータである。

これら二つの表現のうちいずれが「正しい」式であるかは議論の分れるところである。ヴィッカリーは、文章から得た式の方が実際のデータに、曲線の全域にわたって、よく一致することを示し、さらに、この場合はゾーンをいくつに区切っても曲線の形は変わらないこと、すなわち核を明確に定める基準は何もないことを証明した。⁶⁵⁾ これに対して、ウィルキンソンは、グラフによる表現から得た式の方が実際のデータによく適合する、と主張している。⁶⁶⁾

一般には、これら二つの相反する主張の間に立って、いずれが実際のデータによりよく適合するかをさらに多くのデータによって「定める」べきである、と考えるのが普通であろう。モデルの優劣は実際のデータへの適合

度によって決められるからである。しかしながら、確率現象の数学モデルを構成する場合、いかなる対象も同じモデルで記述できるのが最良かもしれないが、このことは必ずしもつねに要求される条件ではない。どのデータもある程度のバラツキを持ち、またその近似であるモデルは、そこに含まれるパラメータによるあてはめの自由度を持つため、モデルの選択は一般に困難な作業である。さらに、いずれの式を後の分析に利用するかは、対象分野、実験目的、データの近似度などによって異なるであろう。ブラッドフォードの法則を図書の管理運営に応用する場合を考えると、大体のめやすとしては、 n が小なる部分に注目する（すなわち、小図書館の蔵書構成を考える）場合には（6）式、 n が大なる部分に注目する場合には（5）式の方が良いであろうといえる。

2) 文献の特性を表わすパラメータ

どのような主題の書誌についても、さらに次節で述べる実験方法のいずれによっても、雑誌とその利用の指標との関係は似たような曲線で表わされるので、この曲線あるいはその数学的表現から1,2個のパラメータを抽出し、これによってある分野の文献の特性を示すことができれば、その分野の概観を得、将来の発展を予測し、蔵書構成を考え、あるいは書誌を編纂・評価したりするのに便利であろう。現在、厳密な理論的基礎づけのなされた、そのようなパラメータは得られていないが、いくつかの提案はある。

その第一は、現在ブラッドフォード乗数 (Bradford multiplier) と呼ばれているパラメータで、ブラッドフォードが用いた核とその次のゾーンの雑誌数の比 n である（(1)式を見よ）。(n が大きいことは分散の程度が大きいことを意味する。しかし、この n は核の大きさと結びついており、核を特定することができれば、分散の様子を示す有効なパラメータとなりうるが、現在、核を一意的に定める基準はない。

第二は、コールの「文献分散係数」(reference scattering coefficient) である。⁶⁷⁾ 彼は、 $R(n)/R(N)$ 対 $\log(n/N)$ のグラフを描いた場合、その直線部は

$$\frac{R(n)}{R(N)} = 1 + \alpha \log \frac{n}{N} \quad (7)$$

で表わされ、この傾き α がいろいろな分野によって異なることから、これで文献の特性を表わしうると考えた。関連文献が多くの雑誌に分散していればいかに程 α は小さ

くなるからである。ところが、 α は雑誌の総数と簡単な関係にあることがブルックスの得た式を用いて証明することができるので、「文献分散係数」はそれ程有用性の高いパラメータであるとは思われない。

文献と雑誌との関係から分散を表わすパラメータを得ようとするこのような試みに対して、第三のパラメータとしてゴフマンらの提案になる一著者あたりの論文数 g がある。⁶⁸⁾ 彼らは、この g が文献の特性をあらわすパラメータとなりうる根拠として、彼らが調査を行なった「住血吸虫病」(schistosomiasis) および「肥胖細胞」(mast cell) の二つの性格の異なる研究分野の文献について、雑誌総数の比、著者数の比、一つの関連論文しか含まぬ雑誌数の比はすべて似たような値をとるのに対して、 g は異なる値をとり、またこの g はいずれの分野についても過去 80 年にわたってほとんど一定であること、および、 g はゾーンを最も細かく分割したときのブラッドフォード乗数とほぼ一致すること、などをあげている。(なお、この場合、同数の関連論文を含む雑誌が別々のゾーンに分配されることもあるが、このことは問題ではない。)

これらのパラメータのうちでいずれが最も有効なパラメータであるかはわかっていない。それを決めるにはブラッドフォードの法則を説明できるような確率モデルをつくるが必要とされるからである。そのような確率モデルの中では、ケンドールのモデル⁶⁹⁾ が最も洗練されたものであるが、これも曲線の全体を説明できるものではない。

3. 実験方法

1) 各種の実験方法

ブラッドフォードの法則に関する実験は、一般に、次のような手順で行なわれる。

- (1) 実験方法の選択
- (2) 対象の選択
- (3) 計 数

雑誌のタイトル毎にカードをつくり、ある雑誌に掲載された論文が現われる度に該当するカードに記入していく。⁷⁰⁾

(4) リストの作成

カードを関連論文の多い順に並べ、雑誌数と関連論文数との関係を示すリストを作成する。

(5) 若干の計算

論文、雑誌の各々について、雑誌の生産性の

高い順に累積和をとり、これをそれぞれ $R(n)$, n とする。

(6) グラフ化

$R(n)$ 対 $\ln n$ の関係を片対数グラフ上にプロットする。

さて、「応用地球物理学」および「潤滑」という二つの主題について書誌を作成し、これに収録された論文とその掲載誌との間に成り立つ関係として定式化されたブラッドフォードの法則は、現在、引用文献調査で行なわれたタイトルの分散に関する研究、あるいは図書館における雑誌の利用度などと同一視されて、雑誌と雑誌の重要度を示す何らかの指標との間に成り立つ関係として一般化されている。このため、さまざまな方法が提案され、それに基づいて実験が行なわれている。計数以後の作業は単純であるのに対して、どの方法によるかは重要な問題であるので、ここでは主な実験方法について、それぞれの方法のもつ特長・欠点などにふれながら、考察することにして。

(1) 引用文献による方法

(a) 一次雑誌を利用する方法

一次雑誌に掲載されている論文の参考文献の掲載誌を調べる方法。これについては 2) で詳しく述べる。

(b) 引用索引を利用する方法

Science Citation Index (SCI) のような引用索引を用いる。本質的には一次雑誌による方法と変わりはないが、計算機を使って計数を行なう場合はもちろんのこと、手作業による場合でも作業はずっと楽になる。マーチンらはこの方法によって英国で発行されているすべての科学雑誌の評価を行なった。⁷¹⁾

(c) レビュー誌・レビュー論文を利用する方法

一次雑誌ではなく、レビュー論文にあげられている参考文献の掲載誌を調べる方法である。この方法は重要な論文のみを取り出すことができるという特長を有するが、当然のことながら、この選択およびカバーする範囲は執筆者の判断によって決められるので、この特長は同時に欠点ともなるおそれがある。また、すべてのレビュー誌が年一回以上発行されるとは限らないこと、および年刊の場合でも毎年同じテーマがとりあげられるとは限らないことなどにより、大きなサンプルが得られない、あるいはサンプルに偏りが生ずるなどの欠点をもつ。セングプタは、この方法を「客観的な分析方法」としてし

て勧め、生化学⁷²⁾ および医学⁷³⁾ の二つの分野について実験を行なった。

(2) 書誌による方法

(a) 抄録誌・索引誌を利用する方法

ある学問分野の概観を得るには便利な方法であるが、一般に、コア雑誌の論文が必要以上に重視され、周辺領域を完全にカバーしているとは必ずしもいえない。さらに、ある種の一般索引誌では、雑誌が主として扱っている分野によって自動的に論文の分類が行なわれている。ある分野の文献の分散の様子を知るために抄録・索引誌を利用するよりは、むしろ逆に、この収録範囲の完全性を確認するためにビブリオメトリクスの手法が用いられる。

(b) 主題書誌を利用する方法

包括的である与否にと拘わらず、できの良さは編纂者の資質によって決まる。編纂者が重要な論文のみを選択して作成した書誌による方法は、レビュー誌による方法と同じ特徴を持つ。主題書誌による場合も、文献の分散を調べるために特に書誌を編纂する場合⁷⁴⁾ と、他人の編纂した書誌を利用する場合⁷⁵⁾ とがある。ブラッドフォードの最大の貢献は、文献分散の定量的関係を我々に教えてくれたことと並んで、完璧な包括性を目ざして主題書誌を編纂し、これによって完全な書誌が持つと思われる曲線のパターンを得たことにある。

なお、図書、モノグラフの重要度を知るためには論文・単行書の参考文献を利用することができるが、文献の形態の違いによる利用の性格の違いが明らかにならないと、雑誌との相対的重要度の比較は難かしいと思われる。

(3) 図書館の利用統計による方法

図書館でよく利用される雑誌ほど研究者にとって重要であるので、利用統計は重要な雑誌を識別するのに使うことができる。さらに、実際に貸出された回数ではなく、貸出し要求あるいは複写申込みの回数も、同じ理由で、分散の様子を知るための基礎データとなりうる。ただし、「分散」の様子を知るためには、分野の特定と大きなサンプルによるバラツキの除去という矛盾した要求を満たさなければならない。同一分野の研究を行なっている大小二つの機関があるときは、大きな機関の図書館の利用統計に基づいて、小さな機関の図書館は購入雑誌の決定なを行なうことができる。

ところで、ある雑誌が図書館に要求された、あるいは図書館から借り出されたことは、必ずしも実際に利用されたことを意味しない。なぜなら、マーチンらも述べているように、“ある図書館から借り出された論文は、おそらくこれは索引誌やカレントアウェアネスサービスによって論文のタイトルを知った結果借り出すのであろうが、読んでみると利用者の感じていたニーズとは無関係であることがわかり、その結果、借出者には利用されないこともあるので、貸出雑誌のリストは雑誌の真の利用率とは正確に対応していない”⁷⁶⁾ からである。

2) 一次雑誌の引用文献による方法の特長と欠点

上述した諸方法のうちで、現在最も多く利用されているのは、一次雑誌中の論文の引用文献を調べる方法である。これは、グロスらが、購入雑誌選定のために行なった調査で、用いた方法でもあり、最も古くから用いられている方法である。このため、この方法について少し詳しく検討してみよう。

この方法の特長および欠点として、ブリッテンらは、次のような点をあげている。⁷⁷⁾

特 長

- (1) 最も多く引用される雑誌は最も多く要求される雑誌となる。
- (2) 引用によってある質の評価がなされている。雑誌のサイズの効果を計算に入れることができる。
- (3) 主題・言語などについての分析をすることもできる。
- (4) 利用度減少のパターンも知ることができる。
- (5) 引用文献の共通性によって対象雑誌を分類することができる。⁷⁸⁾
- (6) 利用の実態(例えば、著者の自己引用)が得られる。⁷⁹⁾
- (7) 引用された雑誌のタイトルのみでなく、引用された論文の分析も行なえる。

欠 点

- (1) 引用の実態と研究者の文献利用との関連が充分良くわかっていない。⁸⁰⁾
- (2) 最も頻繁に引用される文献は最も手に入りやすい文献であるという傾向がある。重要な資料も図書館で利用できないために引用されないこともある。⁸¹⁾
- (3) 少数の雑誌を使う場合その選択が困難である。
- (4) 調査雑誌中の論文はその雑誌に発表された論文

を多く引用する傾向がある。

- (5) 調査雑誌を多くすると、作業量は尨大になる。
- (6) ある主題の文献の「全体的展望」を与えない。⁸²⁾
- (7) 雑誌以外の資料を調査対象とする場合、選択、サンプルの大きさなどの点で、困難な問題が生ずる。
- (8) 一国で発行された雑誌のみを調査対象とする、大きな偏りが生ずる。

ここにあげられた特長・欠点のいくつかは注で検討しているので、さらに、引用文献調査に対するブロードマンの批判を紹介しておこう。その批判の骨子は次の通りである。⁸³⁾

まず、引用文献による方法は次のことを仮定している。

- (1) ある定期刊行物の専門的研究者に対する価値は、専門文献に引用される回数に比例している。
- (2) 作表のベースとして使用する1あるいは2以上の雑誌はその分野全体の代表である。
- (3) 2以上の雑誌をベースとして使用する場合、それらはすべて等しい重みを持つ。

この仮定(1)が成立しているか否かを検証するため、コロンビア大学内科外科カレッジ生理学科のメンバーの投票によって決められた雑誌の「価値」の順位(これをAとする)、レビュー誌、一次雑誌2誌の引用文献によって決められた雑誌の順位(それぞれB、Cとする)のそれぞれの間の相関係数が算出された。⁸⁴⁾ その結果、B・C間の相関は認められたのに対して、A・B間、A・C間ともに若干の傾向が認められたにすぎなかった。また仮定(2)、(3)についてはレビュー誌1誌、一次雑誌4誌のそれぞれの引用文献により決められた順位の間に相関の認められない組合せが多くあった。これより、上の三つの仮定はすべて成り立っていない、との結論を下し、「これ(グロスらの方法)は、信用できない結果を与えるものであると同時に、いく分非科学的かつ非学問的であるように思える、⁸⁵⁾とまでいっている。

これらの批判に対していくつかの疑問が生ずる。まず最初の批判に対する第一の疑問は、科学者はそれほど容易に雑誌の順位づけができるのであろうか、という疑問である。いくつかのテーマを持っており、テーマによって重要な雑誌が異なる場合はとくにこのことがいえる。また関心対象は次第に変化していくものである。第二の疑問は、これらのすべての票に同じ重みを与えて良いものだろうか、という疑問である。多くの調査で明らかにように、科学者の間には、雑誌の利用度、論文の生産量

において非常に大きな違いがある。第二の批判に対しては、2以上の雑誌を調査対象とする方法は、1誌にたよる場合に生ずる偏りを修正するために考えだされた改善策であり、調査雑誌間の相関が小さいことは、複数の雑誌を利用する根拠とはなりえても、この方法の有効性を減ずる証拠とはなりえない、ということがいえる。また雑誌のサイズの効果を補正する方法はある。

このように、雑誌自体の重要さは研究者によって異なり、また重要な雑誌はとりくんでいるテーマに従って変わっていくものであるから、とくに機関の関心対象が個人の関心対象の総和として決められる図書館では、「真の価値」の尺度づけは困難である。極端に言えば、ブロードマンが行なった調査に対する回答の総和が生理学の分野全体を反映しているという保証は何もない。

3) 実験方法の選択

1) にあげられた各種の実験方法により得られた結果の間には大きな違いがあり、⁸⁶⁾ このため、雑誌選択の方法として、いずれが優れた方法であるかについて定まった意見はない。一方では、自己の採用した方法が最も優れた方法であると考えて、この観点からのみ、それ以外の方法が批判され、他方では、各方法による結果の差異のみが強調される。しかし、これらの議論は、暗示的には科学者が実際に利用する度数の高い雑誌によって図書館の蔵書構成はなされるべきであるとの前提に立っているようではあるが、筆者には、雑誌選択の基準をあいまいにしたまま行なわれている、きわめて実り少ない議論であるように思われる。図書館はいかなる目的に最も適う雑誌をそなえるべきかを不明確にしたまま、雑誌の選択方法の優劣を論じることとはできない。そして、すべての図書館が同じ目的を持っているとは必ずしもいえない。このため、図書館がいかなる雑誌をそなえるべきかを定める基準は図書館によって異なる。

そのような基準として次のようなものが考えられる。

- (1) 雑誌の利用度数
- (2) 機関の活動への貢献度
- (3) 論文執筆への貢献度
- (4) 雑誌の重要度

したがって、実験方法の選択の問題はよるべき基準の決定とその基準を知るのに最も適した実験方法は何か、実験方法のもつ誤差、それを補正する方法の順序で議論が進められるべきである。

利用度数には、図書館統計で使われる利用度数と、ある雑誌中の論文が実際に読まれる度数とがある。後者

は、貸出し後、放置され読まれずに返却される雑誌や、論文のタイトルから判断した内容が実際の内容と違っていたため利用されずに返却される雑誌があるので、前者とは異なるものであるが、図書館の貸出し統計、複写要求の記録などは利用度数の近似値を得るために利用できる。

機関の活動への貢献度は、活動の成果を評価する基準が定められない限り、これを測定することはできない。もし、ある機関の目的が科学研究を行なうことであるなら、その成果は一般に論文として発表されるので、これに対する雑誌の貢献度を測る基準は(3)とほぼ同じになる。そして論文執筆への貢献度は引用文献調査によってその近似値を得ることができる。

重要度とは、ある雑誌に含まれる関連論文の質と数とに基づいて決められる尺度である。この尺度は利用の目的によって異なり、また利用者のバックグラウンド、経験年数、アプローチの方法などによっても違いの生ずる主観的な量である。これは利用者へのアンケートなどによって知ることができる。

これらのうち、いずれの尺度によるかという問題と並んで、この尺度をどの程度の規模ではかるかは重要な問題となる。全世界の普遍的な利用を知るために行なわれた調査結果に基づいて小図書館の蔵書構成を行なうことはできない。

4. 文献の寿命と成長

論文の年令を t 、年令 t の論文の数を $p(t)$ とすると、文献の年令分布は次式で表わせることが知られている。⁸⁷⁾

$$p(t) = p(0)a^t, \quad (0 < a < 1) \quad (8)$$

もしくは、

$$p(t) = p(0)e^{-\lambda t} \quad (\lambda > 0) \quad (9)$$

または、最近 t 年間に発表された(すなわち、 t より「若い」)文献の割合 $p(\leq t)$ を求めて、

$$p(\leq t) = 1 - e^{-\lambda t} \quad (10)$$

文献の寿命を示す尺度としては、バートンらが核物理学における放射性同位元素の崩壊とのアナロジーから「半減期」(half-life)を提案して以来、これが良く使われる。⁸⁸⁾ ある年令 τ より若い文献の数が文献全体のちょうど半分になるとき、この τ を「半減期」という。このとき τ は、式(9)あるいは(10)の λ を使って、

$$\tau = \frac{\ln 2}{\lambda}$$

となる。

文献の半減期に関しては各分野につき多くの論文が発表されており、これらの結果から、基礎科学は応用科学より長い半減期を持つこと、基礎科学の中でも数学・生物学などに比べると物理学の半減期ははるかに短いことなど、多くの事実が知られており、得られた数値の間には調査者によってある程度のバラツキはあるが、大体の傾向はわかっている。これに対して、バーンは、これまでに発表された結果の間には大きなバラツキがあり、これらを総合すると各分野の年令分布曲線は重なってしまい、そのため、明確に識別しうる半減期などというものは定義できず、「ぼやけた半減期」('fuzzy' half-life)しかないのだと言っている。⁸⁹⁾ しかし、1899年から1962年までの利用パターンをすべて含めることによって、時代の変化による年令分布の変化をまったく無視し、さらに平均値の差の検定などを行うことなしに、年令分布曲線の端部のわずかの重なりによってこのような結論を下すことは誤りであると考えられる。利用度減少とは、文献の利用が年令の単調減少函数になるという事実を指すと考えるなら、あらゆる分野の文献について利用度減少は存在する。そして、これは、月の単位まで厳密に定めうる、正確な尺度としては扱えなくとも、その分野の文献の特性を示す指標としては充分意味がある。

しかしこのことは、必ずしも、ある文献が本来持っている価値が年月の経過とともに減じていくことを意味しているわけではない。利用度減少における利用度とは、1論文あたりの利用度ではなく、ある年に発表された論文の延利用回数の意味だからである。これより、もし文献の増加率が利用度の増加率より大きければ、利用度減少といった事実はなく、1論文あたりの相対的価値は昔の論文ほど高いということになる。これは、現在、成長(growth)の問題として利用度減少と関連づけて議論されているが、この問題は、ゴフアードらにより、ワイスの論文⁹⁰⁾に対する批判として、初めてとりあげられた。ゴフアードらは、「もしワイスの曲線に実際の論文数による修正をほどこしたなら、それらは確実に横ばい状態になったであろうし、さらにその傾きが逆にさえたかもしれない」⁹¹⁾と述べて、論文の相対的価値は年令によって変わらないものであると主張した。⁹²⁾

しかしながら、実際の利用度減少は、文献の成長と論文の相対的価値の減少の両者によって生じていると考えられる。プライスは、*Science Citation Index*より得た科学における引用文献の年令分布曲線から、利用度減少

の「70%はそれまでに発表されたすべての科学論文がその年令によらずランダムに分布していることによっており、30%は最近の文献への高度に選択的な参照である」⁹³⁾との結果を得ている。

このため、いろいろな要因を考慮に入れ、これによって修正を加えた、いくつかの半減期がラインによって提案された。⁹⁴⁾ 彼は「半減期」には次の4種類があり、これらは明確に区別されねばならないと述べている。

(1) 見かけの半減期 (apparent half-life, or, median citation age)

引用文献調査において引用文献の半数がその間に現われている期間。(これまでわれわれが半減期と呼んでいたものと同じ)。

(2) 修正された半減期 (corrected half-life)

見かけの半減期から成長の要素を除去することによって推定された半減期。

(3) 真の半減期 (true half-life)

ある文献の集合を構成している雑誌の総利用回数の半分の数だけの利用がその間に行なわれた、あるいは行なわれることが期待されている(現実の、あるいは期待)期間。

(4) 雑誌別半減期 (item half-life)

個々の雑誌の総利用回数の半分の数だけの利用がその間に行なわれた、あるいは行なわれることが期待されている(現実の、あるいは期待)期間。

これらの半減期のうちで、図書館員が知る必要があるのは、「個々の雑誌をどのくらいの期間所蔵しておく必要があるか」であるから、このためには「雑誌別半減期」を知ることが望ましい。しかし、これについては実際の場合に適用できるようなデータは得られないので、ラインは、「修正された半減期」を雑誌別半減期の平均値である真の半減期の粗い近似として用い、これに基づいて雑誌の廃棄などを行なうことを提案した。⁹⁵⁾

これに対してヴィッカーは、利用度減少には、現実には、文献の成長・相対的価値の低下のいずれの要因も効いているので、この二つの要因を抽出したことの意義は認めるが、実務上の決定は「見かけの半減期」に基づくべきである。⁹⁶⁾ とした。

しかし、ある雑誌の有用性が、ある分野の関連論文をいくつか含むか、あるいはその雑誌がどれくらい利用されたか、によって決められるのなら、実務上の決定は「雑誌別半減期」に基づいて行なわれるのが当然ではないだろうか。そして、システム分析の観点からすれば、まず

第一に知るべきことは雑誌の「平均的な」特性であるから、「真の半減期」が重要なパラメータとなろう。これに基づいた分析はVで行なわれる。そこでは、ある分野の個々の雑誌の年令分布は、その分野の雑誌全体の年令分布曲線と同じパラメータを持つ指数分布に従う、と仮定される。

5. いくつかの注意

ブラットフォードの法則を実際の図書館運営のさまざまな局面に応用する場合、いくつかの不十分な点はあるが、この法則が有効な手段であることは、これまで述べてきたことから、明らかであろう。しかし、この法則は、無制限に適用できる万能薬ではない。この法則自身がある限界をもち、また「実験」をする際に、およびその結果に基づいて何らかの決定を下す際に、注意すべき点が少なからずあるからである。

第一に、実験の手順から明らかなように、この法則は、関連論文の量のみを扱うもので、その質は一切考慮に入っていないという限界をもっている。同一テーマの論文でもその質には大きな違いのあることは周知の事実であり、ある程度の選択機構を媒介としている引用文献やレビュー誌を利用することの意義もここにあるわけである。また、コア雑誌の論文は主として研究論文であり、周辺雑誌の論文には解説論文が多い傾向にあるので、コア雑誌の論文の方が一般に重要であるともいえるが、狭い分野を対象としている小図書館以外では、研究論文と解説論文といった二分法すら余り意味のあるものではないし、そのほかに考慮すべき要因も多い。しかし、論文の質を、ビブリオメトリクスの手法を適用して、判断する方法もいくつかある。

第二に、この法則の適用対象となる図書館の専門分野を特定することは困難な問題である。これまでに発表された研究結果に基づく雑誌の重要度の判定は、広い領域を対象としたものであって、一般の専門図書館ではこの判定にそのまま従うことはできない。さらに重要なことは、この問題は、例えば半導体工学といった、狭い領域の文献の特性が明らかにされれば解決される問題ではない。なぜなら、例えば、ある製造企業の図書館を考えた場合、どの分野の文献をその図書館で収集すべきかは現在の学問の分類体系の中に明確な境界を画して位置づけられるものではないし、また、同種の企業でもまったく同じということはいえないからである。このことは、一人一人が異なる専攻分野を持っている大学図書館において、とくに問題となる。

第三に、ある分野のすべての関連論文が数十タイトルあるいは数百タイトルの雑誌から得られることは明らかであるが、コレクションの完全性を求めるのは、一図書館においては、賢明な方針とはいえない。周辺部分、とくに関連論文を一篇しか含まぬ雑誌のタイトルは大幅に変化していくのが普通であるから、世界中の雑誌すべてを収集しない限り、完全なコレクションは望めない。ただし、コア雑誌以外に掲載された利用度の高い論文の別刷あるいはコピーをそなえることは、これとは別問題である。

第四に、第三とも関連する問題であるが、確立された分野においても、関連論文を掲載している雑誌のタイトルは年とともに変化していくものであり、適切なコレクションを維持するためには、最低5年に1度くらいの継続的な調査が必要である。

第五に、雑誌の価格には違いがあるので、相対的な利用度のみで購入するか否かを判断するのは誤りである。関連論文数1の差で偶々選定リストに加えられた雑誌でも、高価なためそれを購入しなければ、合計するとその数倍の関連論文が得られる数種の雑誌を代わりに購入できることもある。

第六に、ある分野の調査結果に基づいてその分野に属する図書館がすべて同じ雑誌を購入しても、より多くの雑誌への接近性という観点から見ると、図書館の有効性は増大するどころか、「科学的方法」を導入する前より、かえって減少する結果になる。このため、一図書館の計画ではなく、より広い視野に立った蔵書構成計画が必要である。この法則はそのための方法をも提供する。

第七に、各種の実験方法によって得た結果は、IV. 3. 3) で述べた雑誌選択基準を得るための近似値でしかない。貸出し統計による場合はブラウジングの効果などが、引用文献による場合は引用に現われない文献の貢献度などが明らかにされてはじめて、各種の実験方法は信頼性の高い方法となりうる。

その他、特定の機関内の利用者の多くにはなじみのうすい外国語で書かれた論文で占められている雑誌、購入が困難な雑誌、発行から受入れまでの遅れがひどい雑誌なども他と同格には扱えないであろう。

V. ビブリオメトリクスの応用

これまでに述べてきたのは、統計的書誌学およびブラッドフォードの法則の実験手順をみると明らかのように、計数という単純な方法によって知ることのできる科

学者と科学文献の特性のみであった。もちろん、初歩的な方法が初歩的な知見しか与えないとは必ずしもいえるわけではなく、論文の計数のみによって明らかにしうる特性が多くあることは上に見た通りであるが、ビブリオメトリクスの諸方法は、論文あるいは雑誌のような固定された単位を越えて、よりミクロな分析あるいはマクロな分析が記録された知識に加えられたとき、そして、それらがより厳密な数学的方法と結びついたとき、より広い適用範囲を持つことができる。たとえば、ブルックスは「現在における書誌の数量的分析の主要な目的」として次の5点を挙げている。

- (1) より経済的な情報システム・ネットワークの設計
- (2) 情報処理プロセスの効率の改善
- (3) 現在の書誌サービスの欠陥の識別と測定
- (4) 出版傾向の予測
- (5) 情報学の理論を発展させる基礎を形成しうであろう経験則の発見と解明⁹⁷⁾

しかし、ここにあげられた諸目的は、ある程度抽象的に表現されているため、それなりに広い適用範囲を示しうるものであるが、やはり、情報学の基礎理論と図書館・情報システムの効率的な運営方法の模索という図書館学研究者のこの分野の研究動機をそのまま反映したものでしかない。いくつかの科学的知識の相関が引用で示されること、ある特定のテーマに関する研究の盛衰が文献量から知ることができること、などを考えただけでも、ビブリオメトリクスの応用面は上記の諸目的にとどまるものではなく、広く「知識社会学」、「科学の社会学」、「科学の科学」にまで及ぶものである。ここでは、便宜的に理論的应用と実際的应用とに分けて、応用例のいくつかをとりあげてみたい。

1. 理論的应用

1) 文献の分類

引用は原論文と引用論文との間に何らかの科学的知識の相関関係があることを示しているの、同じ論文を引用している二つの論文の間にも何らかの共通性があると考えられる。その共通性には、研究方法、研究対象などがあるが、これらをすべて含むものとして、若干の例外は覚悟の上で、同じ論文を引用している二つの論文はともに同じテーマを扱っていると考えるとそれほど大きな間違いをおかすことはないと考えられる。このような考えに基づいて、引用文献の共通性によって文献の分類を行

なうことができる。一般の検索語と並んで、引用文献を指定することによって検索効率を高めるといった利用法もあるが、⁹⁸⁾ ここでは、引用文献のみに基づいた科学文献の分類に話を限る。

引用文献の共通性による文献分類の方法はケスラーによって開発された。彼は、⁹⁹⁾「専門論文の引用文献は、著者が自分の活動している知的環境を示す一つの方法であり、もし二つの論文が同じ引用文献をあげているなら、これらの論文の間にはある関係の存在することが示唆されている」という仮説に基づいて、多数の論文を小さなグループに分ける方法を考えた。分類は次のようにして定められる論文間の結合強度 (coupling strength) の大小によって行なわれるので、彼はこの方法を「書誌的結合」(Bibliographic Coupling) による方法と呼んでいる。

- (1) 二つの論文のいずれにも引用されている文献が一つある場合、これら二論文間の結合単位を 1 とする。
- (2) 所与の試験論文 P_0 との結合単位が少なくとも 1 の論文の集まりは関連グループ G_A を構成する。
- (3) G_A 中の論文と P_0 との結合強度はこれらの間の結合単位の数 n で測られる。¹⁰⁰⁾

彼は、*Physical Review* の一卷中の個々の論文を P_0 とし、他の論文との結合強度を求めて、この一卷中の論文を分類し、さらに、この結果をこの雑誌の編集者が用いている Analytic Subject Index (ASI) によって分類した結果と比較して、これら二つの結果の間には高い相関のあること、すなわち、熟練者が論文を読んで行なった分類とほとんど変わらない結果を得たことを報告している。

この方法の特長として彼は次のような点をあげている。¹⁰¹⁾

- (1) 言語とは無関係。すべての処理は数のみで行なえる。
- (2) 論文を読んだり判定したりする必要はない。本文さえもいらぬ。
- (3) ある論文に静的な分類番号あるいは恒久的な索引番号をつけない。グループ分けは学界のカレントな利用・興味を反映して変えうる。

「書誌的結合」の一変形と考えられるのは、スモールの「同時引用」(co-citation) による方法である。¹⁰²⁾ これは、二つの論文をともに引用している論文の数によっ

て、それら 2 論文間の関連の度合を測るものである。彼はこの方法の、「書誌的結合」の方法に対する、優越性を主張しているが、それは他分野における研究とのアナロジーによって新しい法則などを得た場合、これらの関連は、「同時引用」では示されるが、「書誌的結合」では示されないという点に限られる。

2) 科学雑誌・学問領域の特性

i) 科学雑誌の特性

雑誌の類型化を行なう最も簡単な方法は、引用文献の寿命と 1 論文あたりの引用文献数による方法であろう。プライスは、引用文献のうち最近 5 年以内に発表された論文の割合 (これを「プライス指数」(Price's index) と彼は名づけている) を横軸にとり、1 論文あたりの平均引用文献数を縦軸にとって、延べ 154 の雑誌を二次元座標上の点として表わして、雑誌の類型化を試みた。¹⁰³⁾ 彼の定義により、横軸で左側にくる (プライス指数の小さな) 雑誌ほど「古文書発掘的」(archival) な雑誌であり、右にくるに従い、「研究の第一線」(research front) にある雑誌となる。また縦軸については、上にいくほど「レビュー誌」の性格が強くなり、下にくると「権威をもって」(*ex cathedra*) 書かれた論文を多く含んでいるといえる。

これらの点のあつまりは、それぞれ 16 (引用文献/1 論文)、32% をメジアンとして、外側にいくに従って少なくなる。しかし、一論文あたりの引用文献数という尺度は、プライス自身も認めているように、真に独創的な論文と評論的な論文との区別が、少なくとも数としてあらわれている限りでは、つけられないために、余り有効な指標ではない。これに対して、プライス指数は、その指数の大小によって、ハードサイエンス、ソフトサイエンス、ノンサイエンスのいずれにその雑誌が属するかを知ることができる。このため、座標上で意味を持つのは横軸上の位置だけとなるので、プライス指数のみによる類型化の結果だけを見ると、このような事実はすでに半減期に関する研究で知られていたことであり、何もこのような指数をつくりだす必要はないことになる。この意味では彼は成功したとはいえないが、ここで重要なことは、雑誌の特性を一つの指標のみで記述することは困難であるが、二つ以上の指標によれば可能になるのではないかと考えられることである。もちろんその前に有効な指標の探究がつけられねばならない。

ii) 学問領域の特性

引用文献の寿命は学問分野の進歩の速さを知るめやす

となるが、それぞれの分野の特性を示す尺度として、他の分野との相関の強さが考えられる。ただし、学問領域の特性を明らかにするためには論文の各々について、その属する分野を決めなければならない。

このような学問領域間の相関関係の強さを示す尺度として、アールらによる“自己引用度”(self-citation)と“自己利用度”(self-derivation)がある。¹⁰⁴⁾ここで“自己引用度”とは、著者あるいは雑誌の自己引用よりはずっと広く、学問分野についていわれており、“ある分野が他の諸分野の文献にどの程度依存しているかを示す指標”と定義されている。また“自己利用度”は“ある分野が他の諸分野にどの程度貢献しているかを示す指標”と定義されている。具体的には、ある特定の分野Dとそれ以外の分野とを考えて、これらの分野の原論文、引用論文(ある論文を引用している論文を原論文、引用されている論文を引用論文と一応しておく)の数がそれぞれ

表 3. 原論文、引用論文の属する分野による引用関係の分類

		原論文の属する分野	
		分野 D	D 以外
引用論文 の属する分野	分野 D	α	β
	D 以外	γ	δ

表3のようになったとする。このとき、分野Dの自己引用度 c は

$$c = \frac{\alpha}{\alpha + \gamma} \times 100 \quad [\%]$$

で示され、 c が 100%に近い程、他の諸分野の文献への依存度は低くなる。また、自己利用度 d は

$$d = \frac{\alpha}{\alpha + \beta} \times 100 \quad [\%]$$

で示され、 d が 100%に近い程、分野Dの他の諸分野への貢献度は低いことになる。実際に調査を行なう場合、対象とする分野は余り狭くとりすぎると論文の属する分野の特定が困難になり、またサンプルも小さくなるので、ある程度の制限はあるが、十進分類法の「綱」(division)以上を単位にとれば、作業はそれほど難しくない。

このような指標は、学問領域間の相関を示しているもので、実務においても、蔵書構成、配架のめやすを得るのに有用である。例えば、アールらは、社会科学、科学、技術の各分野の自己引用度、自己利用度について次のような結果を得ている。

	自己引用度	自己利用度
社会科学	58%	96%
科学	70	78
技術	81	62
科学技術全体	97	90

この結果のみからも次のようなことがいえるであろう。

- (1) 社会科学は広い分野の文献を必要とするので、蔵書は多岐にわたらざるをえない。しかし、その成果は他分野ではほとんど利用されない。
- (2) 科学と技術は密接な関連がある。このことは、蔵書構成・配架のいずれにおいても考慮されねばならない。
- (3) 科学技術全体を考えると、社会科学などの文献をほとんど必要とせず、それゆえ、他分野の雑誌の購入を考える必要は余りない。

アールらは、この他に政治学、経済学など社会科学の諸分野について同様の調査を行なっている。自己引用度、自己利用度はともに、例えば民法→法学→社会科学のように分野を広くすればする程、大きくなっていくが、社会科学の中で特に興味ある数字を示しているのは教育学で、自己引用度、自己利用度はそれぞれ36%、92%となり、教育学は他分野の文献を多く必要とするが、その成果はほとんど教育学の中で利用されるだけである。

3) 科学の進歩の分析と予測

ある分野の文献の成長が指数函数曲線に従っていることは、すでに「文献の寿命と成長」のところで述べた。しかしこのような成長は、新たに生まれ消えていった多くのテーマがつくりあげたものであり、そのテーマの重要性は時代の移り変わりによって大きく変わっていく。このため、より小さな分野あるいは一つの主題を考えると、その文献の成長曲線はもはや指数函数曲線にはならないが、逆に、この事実をあるテーマの重要性の変遷を示す尺度として利用できることは明らかであろう。このためのツールとしては、当然のことながら、書誌が使われる。このような特定の研究テーマに関する研究の盛衰を調べることは、科学史的興味と同時に、現在および将来の研究機関・研究者に対する資源配分の時期・量の決定、教育機関における新しい学科・研究所の設立の計画などを行なう場合に欠かすことのできない作業となるべきである。

i) 情報伝播の疫病モデル

これは、特定の研究者があるテーマについての研究活動を開始した時点と停止した時点を決める必要があるためミクロな分析が要求される点、およびその基礎となる理論が他分野においてすでに確立されている点で、これまでの諸方法とは異なるものである。この研究を行なうためには、あるテーマに関する研究の歴史的な発展を詳細にたどり、一定の厳密な基準に基づいて論文の選択が行なわれた（さもないければ限り網羅的に収集された）書誌がなければならない。これは抄録誌、索引誌などを逐一調べることによっても行なえるが、作業量は尨大になるので、現在までのところは主題書誌を利用した分析しかない。この貢献者はゴフマンである。

疫病の伝染は次のようなモデルで考えられる。ある集団に保菌者 (infectives) が導入され、未感染者 (susceptibles) のうちの何人かが保菌者との接触あるいは中間媒介物により感染した後、時には一定の潜伏期 (latency period) を経て、罹患する。疫病の蔓延は未感染者から保菌者への移行がある閾値を越えたときにおこる。この間に、死亡、免疫獲得、隔離などにより除去される者 (removals) もいる。このような疫病の伝染を記述する決定論的あるいは確率論的な数学モデルは、疫学、生物数学などの分野ですでに開発されていた。ゴフマンはこのうち、数学的な取扱いの容易さから、決定論的モデルによっている。

個人的接触あるいは多様なメディアによるアイデアや情報の伝播過程が、疫病の伝染過程と類似であることは明らかであろう。情報伝播の諸要素は疫病の流行の諸要素と次のように対応している。¹⁰⁵⁾

疫病伝染の要素	情報伝播の要素
病 源 体	アイ デ ア
保 菌 者	論文の著者
未 感 染 者	論文の読者
除 去	死亡・興味の喪失

このようなモデルは、文化の伝播、新しい技術・製品の普及などさまざまなプロセスを分析するのに利用しうるが、ゴフマンの場合は、記録物による新しいアイデアの伝播を、それも研究のみに限って扱っている。それゆえ、彼のモデルによれば、ある新しいアイデアの伝播は次のようになる。あるアイデアを持った研究者は、論文を書くことによって、そのアイデアを伝える人（保菌者）になる。その論文を読んだ人（未感染者）のうち何人かにそのアイデアが伝わり、その研究者たちがまた同じテーマについて論文を書くことによって、そのテーマ

の研究は次第に発展していく。（その論文を読んだ者は誰でもその分野の研究者となりうるわけではないから、保菌者を、例えば、疫学に関する論文を書いたことのある人とする、未感染者は医学に関する論文を書いたことのある人などとなろう。）しかし、この途中で、死亡したり、関心が他のテーマに移ったり、退職したりして、そのテーマの研究をやめる人たち（除去）もいる。さらに、他の分野で疫学に関する論文を書く人もいる。これより、その分野のアクティブな研究者の数の変化はこれら三つの要因によって決められる。ゴフマンは、一たん免疫を獲得した者がまた未感染者に戻る確率は0としている。

このようなモデルをつくって、次のような実験を行なう。ある研究者があるテーマに関する論文を初めて書いた年をその研究者の感染年とし、そのテーマについて最後に論文を書いた年を除去年とする。感染年および除去年の決定をそのテーマに関する論文の著者すべてについて行ない、保菌者数の前年度に対する変化の年次推移をプロットする。この手順に従って実験が行なわれたテーマには、ゴフマンらによる記号論理学¹⁰⁶⁾、肥胖細胞¹⁰⁷⁾、ドナヒューによる情報学¹⁰⁸⁾ などがある。

このようにして得られたアクティブな研究者の数の変化から、あるテーマや分野に関する研究の盛衰を知り、その将来あるいは研究のピーク・衰退期を予測し、さらに、それらの分野が発展するための条件を導びくことができる。そして、ゴフマンとハーモンが記号論理学研究の発展に関する研究で行なったように、伝染の始まりは必ず重要な定理の証明、矛盾の発見などと結びついているので、科学史研究などの有力な方法となりうる。

ゴフマンの方法はあくまで単純化されたモデルによっているのであるが、その限りにおいては、重要な要因の見落としはないし、単純さを犠牲にしてまで含む必要のある要因もないと考えられるので、理論については充分であろう。しかし、これを実際に適用する際に、記号論理学のように長い歴史を持つテーマは「除去」を見分けることは容易であるが、ドナヒューが行なった情報学のように、短かい歴史しか持たない分野に適用するのは危険であり、とくに、研究者がわずかしかない場合は注意しなければならない。なお、図書・モノグラフなどの出版物も含めた方が誤差は少ないであろう。ついでながら、アクティブな研究者の数の変化を求める際に、科学者数の増加を考慮に入れるべきであるとの批判が出されるかもしれないが、この要因を考慮に入れて補正を行な

っても、全体的なパターンは若干変化するが、重要なのは曲線の立上り、立下り、ピークの位置などであるので、このような補正はそれほど意味のあるものとは思われない。

ii) 科学研究のライフ・サイクル

ゴフマンの疫病モデルと比較するとずっと単純であるにもかかわらず有効な方法として、山田らによって用いられた科学研究のライフサイクルを見出すための文献調査の方法がある。¹⁰⁹⁾ この方法はIIで述べた統計的書誌学と本質的な違いはない。彼らは、化学を中心とするノーベル賞クラスの研究テーマを例にとり、それもこれらの基礎研究段階に限って、「科学研究のライフ・サイクル」を求めた。その方法は次の通りである。

- (1) ノーベル化学賞・物理学賞および医学賞の受賞対象となった研究テーマの中から化学と関連の深いテーマを選び、このうちで *Chemical Abstracts* の索引語となっている研究、およびこれらに次ぐ若干の重要な研究について、それぞれの項目の下に記載されている研究論文の数を 2~3 年おきに数えあげる。
- (2) それぞれのテーマについて、年間採録論文数に対するそのテーマの研究論文数の割合を求め、この値の年次変化をグラフに描く。
- (3) この曲線を

$$y = \frac{c}{1 + a(x-b)^2}$$

ここで、 x : 西暦年

y : 論文数比

で近似し、パラメータ a, b, c を最小二乗法により定める。このとき a, b, c はそれぞれ次のような意味を持っている。

a : ライフ・サイクル曲線の立上りの傾斜の大きさと最盛期の長さを示す尺度

b : ライフ・サイクルのピークの時点

c : ピークにおける論文数の比(研究の重要度)

彼らは、この方法をペニシリンなど 43 のテーマに適用して、次のような結果を得ている。ただし、ライフ・タイムは、研究の最盛期がどのくらい続くかを示す指標で、上の曲線の変曲点間の距離(もちろん単位は年)がとられる。また、タイム・ラグは、あるテーマに関する研究が最初に発表されてから、そのテーマの研究がピークに達するまでに経過した年数をさす。

- (1) 研究のピークの分布は、1925~34 年と 1946 年

以降の二つのグループにはっきりと分れている。

その原因としては、この間に第二次世界大戦があったこと、研究の内容が質的に変化していることが考えられる。

- (2) 科学情報の伝播速度がはやくなるにつれてタイム・ラグが単調に減少するであろうという予想に反して、発明・発見年とタイム・ラグとの関係は、ともに勾配がほぼ -1 の直線で近似できる、二つのグループに分れる。
- (3) ライフタイムの平均値は 7.4 年で、第二次大戦前と後のどちらのグループもほぼ同じである。研究の重要度とライフ・タイムとの相関は認められない。これは、研究者が一つのテーマを選択すると、成果の大きさやテーマの客観的重要性にあまり影響されずに、ある期間そのテーマを変えないという傾向があることによる。

iii) 科学発展のマルコフ・モデル

あるテーマに関する研究の進歩がたどる道は決して、科学研究者の総数の増加に見られるような、直線的な繁栄へむかう道ではない。「科学研究のライフ・サイクル」で見たように、多くのテーマは、一つのサイクルを完了すると、ほとんど顧みられなくなる。しかし、50年、100年、あるいは200年とつづいている研究テーマは、どのようなパターンに従って発展してきたのだろうか。その分析・予測は、ゴフマンが情報伝播の疫病モデルで行なったように、アクティブな研究者数の変化を知ることにより行なえるが、これを一般化する方法はないだろうか。ゴフマンらは、科学発展の段階をいくつかの状態に対応させ、科学研究の進歩をこれらの状態間の遷移と考えて、一つの手がかりを与えている。¹¹⁰⁾

科学研究の段階は、その時、いかなる情報が蓄積されているか(あるいは、いないか)によって、次の四つの状態に対応させることができる。

状態 E_1 : 不十分かつ無秩序な情報

E_2 : 不十分であるが秩序立てられた情報

E_3 : 充分であるが無秩序な情報

E_4 : 充分なかつ秩序立てられた情報

このような状態間の遷移はマルコフ過程と考えることができる。(実際に、科学研究がこれらのうちのいずれの状態にあるかを決めるのは困難であるが、ゴフマンらはこれを、「特別の関心をひく、あるいは、重要であると考えられる出版物」および「基本的に重要な新しいアイデアが最初に現われたことを示す」出版物の有無によって

定めている。)そして、ひとたび、状態間の遷移確率が求められると、これから、極限分布を知ることによって、各状態の平均停滞年数がわかり、将来の発展の方向を知ることができる。

しかし、マルコフ過程は、事物の状態・性質の変化を記述するものであり、時間的な変化が性質の変化と結びついていない事物にこれを適用するのは誤りであることに注意したい。例えば、ツンデらは、引用は、科学研究の力点が引用論文の属する分野から原論文の属する分野へ移行したことを示していると考えて、「科学の進歩の予測モデル」をつくり、社会科学の各分野間、および自然科学と社会科学との間における将来の研究の力点の移行変わりを「予測」している。¹¹¹⁾その結果、現在社会科学と自然科学の文献数の比は約3対11であるが、将来は大幅に逆転して約17対3になると予測した。しかし、彼らの方法は仮定からすでに間違っている。このことは、マルコフ連鎖の極限分布の性質から簡単にわかることである。¹¹²⁾

このような結論を彼らが導いたのは、マルコフ過程とは状態遷移を扱う理論であることが理解されていなかったためであろう。科学研究とは、初めに存在する資源が、ある変化を受けて、社会科学になったり自然科学になったりするのではないから、引用は、ある分野の他分野への依存度を測る尺度とはなりえても、これによって、上のような、科学発展の予測はできない。また、情報は何度でも使えるものである。

2. 実際的应用

1) 図書館における雑誌の収集と廃棄

グロスらによるタイトルの分散に関する調査の目的は、最も多く利用される雑誌を収集して、最良のコレクションをつくりあげることにあつた。しかし、どの程度の蔵書をそなえるべきかは、利用者、予算、他の図書館の利用のしやすさなどさまざまな要因を考慮に入れなければならない。そのためには、まず、コレクションの大きさとそれによって満たされる要求との関係を知る必要があろう。

上位 n タイトルの雑誌より得られる関連論文数 $R(n)$ は (5) 式より求められる。ところが、この式のパラメータ k は N とほとんど等しいことが証明でき、¹¹³⁾ 結局、 $R(n)$ は、雑誌の総数 N を使って、次式で表わすことができる。

$$R(n) = N \ln \frac{n}{s} \quad (11)$$

これより、上位 n タイトルの雑誌を、必要なバックナンバーを含めて、すべてそろえたときに得られる関連論文の割合 p は、

$$p = \frac{\ln \frac{n}{s}}{\ln \frac{N}{s}} \quad (12)$$

となる。また n 番目の雑誌中の関連論文数 $r(n)$ は (11) 式を微分して、

$$r(n) = \frac{N}{n}. \quad (13)$$

ところで、すべてのバックナンバーをそろえておくことは、スペースの制限により、あるいは新たに購入する場合は経済的制約により、必ずしも望ましいことはいえない。このため、文献の利用度減少を使って、最近の雑誌のみで蔵書を構成したときに、得られる論文の割合はどのように変化するかを考えてみよう。バックナンバーを購入する問題も古い雑誌を廃棄する問題も同じなので、ここでは、廃棄の問題として考える。

年令 t のある雑誌に含まれるある分野に関する論文の数の ($t=0$ のその雑誌に含まれる論文数に対する) 比を a^t ($a < 1$) とする。¹¹⁴⁾ このとき、廃棄を一律に行なうか選択的に行なうかにより、次のような二つの方針が考えられる。

(1) 年令が t 以上の雑誌はすべて廃棄する。

発行後 t 年経過した雑誌をすべて廃棄すると、その時に、上位 n 誌より得られる文献数 $R_1(n)$ は、

$$\begin{aligned} R_1(n) &= \frac{1+a+\dots+a^{t-1}}{1+a+a^2+\dots} R(n) \\ &= (1-a^t) N \ln \frac{n}{s} \end{aligned} \quad (14)$$

となる。当然のことながら、 $R_1(n)$ は、 $t=0$ で 0 、 $t=\infty$ で $R(n)$ になる。これより、この方針に従った場合に得られる文献の全文献に対する割合 p_1 は、

$$\begin{aligned} p_1 &= (1-a^t) \frac{\ln \frac{n}{s}}{\ln \frac{N}{s}} \\ &= (1-a^t) p \end{aligned} \quad (15)$$

となる。¹¹⁵⁾ この方法は、購入年度のみによって雑誌の廃棄時期が決められるので、業務を行なう上からは便利であるが、コア雑誌も周辺雑誌もすべて同等に扱うので、雑誌の間には利用度に大きな差異が生ずることになる。

これを考慮に入れたのが次に述べる方針である。

(2) 利用度数がある基準以下になった雑誌はすべて廃棄する。

この基準として、年令 t の雑誌 1 巻から得られる論文数が $(1-\alpha)N/M$ になる点が選ばれる。¹¹⁶⁾ これは n 番目の雑誌 1 巻から得られる論文数が、経年により、購入年に M 番目の雑誌 1 巻から得られた論文数と同じになったとき廃棄することを意味している。このとき、個々の雑誌をいつ廃棄すべきかは、

$$\frac{N}{n} a^t = \frac{N}{M} \quad (16)$$

より求められ、この方針に従った場合に、コレクション全体から得られる関連論文数 $R_2(n)$ は、

$$R_2(n) = N \ln \frac{n}{s} - \frac{N}{M} (n-s) \quad (17)$$

となり、この値の全文献 $R(N)$ に対する割合 p_2 は、

$$\begin{aligned} p_2 &= \frac{\ln \frac{n}{s} - \frac{n-s}{M}}{\ln \frac{N}{s}} \\ &= p - \frac{n-s}{M \ln \frac{N}{s}} \end{aligned} \quad (18)$$

より求められる。

なお、必要とされる書架のスペースは、雑誌によりまた個々の巻により違いはあるが、所蔵すべき巻数に比例すると考えると、スペースの比較は所蔵巻数によって行なえる。上の二つの方針のそれぞれの場合について、所蔵すべき雑誌の巻数 C_1 , C_2 を求めると、

$$\begin{aligned} C_1 &= nt \\ C_2 &= \log_a(n!) - n \log_a M \end{aligned} \quad (19)$$

となる。なお、一般に、 $n \gg 1$ なので、 $\log_a(n!)$ の計算にはスターリングの公式を使うことができる。書架のスペースだけが制限されている場合には、多くの関連論文を与えるという意味で、最も効率の良い蔵書構成をするには、雑誌の巻数 C がある値を越えないという制約条件の下で論文の割合 p を最大にするという極値問題を解けばよい。ただし、ここではこのような一般的な取扱いはせずに、具体的な数値をあげて方針 (2) の有効性を示しておこう。

【例】 あるテーマに関する論文とその掲載誌との関係を調べた結果、次のような値が得られたとする。

$$N=600 \quad [\text{vols}]$$

$$s=1.5$$

$$\alpha=0.87 \quad (\text{これは半減期 } \tau \approx 5 [\text{年}] \text{ に相当する})$$

このとき、(11) 式より

$$R(N)=3600$$

となる。ここで、上位 60 誌をすべて収集したとすると、これらの雑誌から得られる関連論文の割合 p は (12) 式より

$$p=0.616$$

さて、この図書館では、スペースが限られているために、バックナンバーをすべて保存しておくことは不可能である。そこで、上の方針 (1), (2) のいずれをとるかによって、得られる論文の割合と所蔵すべき雑誌の巻数はどのように異なるかを求めてみる。

(1) $t=15$, すなわち、購入後 15 年経過した雑誌はすべて廃棄することにとすると、

$$p_1=0.539$$

$$C_1=900 \quad [\text{vols}]$$

(2) $M=200$, すなわち、一巻から得られる関連論文の数が 3 以下になった雑誌はすべて廃棄することにとすると、

$$p_2=0.535$$

$$C_2=545 \quad [\text{vols}].$$

これより、雑誌の順位によって廃棄時期を変え、ほぼ同数の関連論文を得るのに、書架のスペースは約 6 割ですむ。

2) 書誌の完全性

抄録誌を中心とする書誌サービスは、現在多くの機関によって提供されているが、このような書誌が重要な論文を洩れなく記載しているかどうかを確認するには、いくつかの書誌を照合したり、あるいは、偶々知った関連論文がそれらの書誌にとりあげられているか否かを精査するしか方法はない。

ところで、我々は、IV. 1. 2) で、「ビブリオグラフ」の作図法を説明したが、ブラッドフォードの得た曲線をはじめとして、これまでに発表された多くの調査結果からわかるように、このようにして作られたビブリオグラフは、 $\ln n$ の増加に対して $R(n)$ は緩い上昇曲線を描いたのち直線的増加に移っていくという、ほとんど同じバ

ターンを示しているの、これはあらゆる書誌の持つパターンであると考えられる。さらに、このビブリオグラフの直線部の傾きより関連論文を含むと考えられる雑誌の総数 N を求めることができるので ((11) 式を見よ)、この直線部を横座標が N の点まで延長して得た曲線 (ブルックスはこれを「完全ビブリオグラフ」と呼んでいる¹¹⁷⁾) は、完全な書誌を持つであろうパターンと考えられる。我々は書誌の完全性をチェックする一つの方法として、この知識を利用することができる。

しかし、このようにいったからといって、完全な書誌を分析して得た曲線は、科学の統一性といった原理から演繹される絶対的真理のゆえに、すべて同じパターンを示さなければならないなどといっているのではない。完全な書誌を分析すれば「完全ビブリオグラフ」が得られるというのは単なる仮定にすぎないのである。ただ、我々の経験によると、念入りに作成された書誌のほとんどはある一定のパターンを示すので、そのようなパターンから大きくずれている書誌は何らかの洩れがあるのではないかと考える正当な根拠があるのである。もちろん、このような仮定は成り立たない場合もあることが、将来の研究によって、明らかになるかもしれない。例えば、現在いくつかの調査結果の n 大の部分で見られる直線のたるみ (droop) は周辺雑誌の探索の不充分さによるものとされているが、これは研究テーマの特殊性を説明するものであることが明らかになる可能性もある。しかし、それが明らかにならない限り、我々は厳しい方の評価基準をとるべきである。書誌サービス機関の側から考えると、これは完全な書誌を作成するための指針となりうる。なお、書誌の完全性のチェックの場合と同じ方法でかつ同じ理由により、上に述べた知識は文献利用調査の欠陥の識別にも利用できる。

VI. む す び

本稿は、これまでの研究成果を展望することを通じて、ビブリオメトリクスの可能性と限界を明らかにする目的で書かれたが、この目的が達成されたか否かは疑問である。研究範囲をある程度示すことはできたが、ブラッドフォードの法則に関しても、各種の実験方法の厳密な比較、サンプリングによる実験の検討、OR 学者が開発した雑誌選択の数学モデルとの比較など残された課題は多い。また、ビブリオメトリクスの応用範囲も本稿で述べた項目のみに尽きるものではない。

しかし、科学文献の統計的分析はいくつかの分野で行

なわれてきたが、ビブリオメトリクスという名称は専ら図書館学研究者の間で使われているために、その研究範囲はおのずから制限される傾向がある。

その例は、ビブリオメトリクスの研究のほとんどが、雑誌選択の基準を得るためにのみ行なわれていることにみられる。たしかに、雑誌選択のために行なわれた引用文献調査は長い歴史を持ち、多くの研究結果が得られている。また、近年、ブラッドフォードの法則以外による方法もいくつか考え出されている。それゆえ、これらの諸方法を適用して雑誌選択の問題を扱ったドナヒューの著書『科学文献の理解—ビブリオメトリック・アプローチ』¹¹⁸⁾ に対するフェアソンの評価(“本書が出版されたことに対して私は大へん満足している。なぜなら、これは……ビブリオメトリクスの揺籃期が終わったことを示しているからである、”(同書序文))もある程度納得できるであろう。しかし、実験結果の機械的な適用のみによって有効な活動はできない。図書館が真に有効な活動をするためには、科学、科学者、科学文献の特性を知ることが必要である。ビブリオメトリクスはこれを解明する有効な方法となりうる。

しかし、ビブリオメトリクスが広い応用範囲をもつからといって、これを一つの独立した分野と考えるのは誤りである。ビブリオメトリクスの方法は単純であるし、他のテーマと比べて、費用、労力もそれ程必要とされないので多くの研究が行われてきたが、大部分が計数処理のみで終わっていることは、このような考え方が研究者の間に広まっていることを示している。しかし、ビブリオメトリクスは1つの手法にすぎず、これを応用する場合には他分野の知識が要求される。たとえば、雑誌選択の問題ではシステム分析の手法や考え方が導入されなければならない、科学の特性の解明には研究組織や科学史の知識が必要とされる。

このような研究は一朝一夕にできあがるものではない。本稿がこのような研究の出発点となりうれば、筆者にとって望外の喜びである。

- 1) Pritchard, Alan. "Statistical bibliography or bibliometrics (Documentation note)," *Journal of documentation*, vol. 25, Dec. 1969, p. 348.

彼はビブリオメトリクスについては単に「図書館および他のコミュニケーションメディアへの数学的・統計的手法の応用」としか述べておらず、この定義は統計的書誌学の定義であるが、名称の変更による研究内容の変更は意図されていないの

- で、より具体的な方をとった。
- 2) Shera, Jesse H. "Of librarianship, documentation and information science," *Unesco bulletin for libraries*, vol. 22, Mar.-Apr. 1968, p. 64. Rep. in Elias, Arthur W., ed. *Key papers in information science*. Washington, D. C., American Society for Information Science, 1971, p. 4-11. Also, in Shera, Jesse H. *Knowing books and men; Knowing computers, too*. Littleton, Colo., Libraries Unlimited, Inc., 1973, p. 269-78.
 - 3) Yelland, M. Research in librarianship. <Whatley, H. A., ed. *British librarianship and information science, 1966-1970*. London, Library Association, 1972> p. 309.
 - 4) 記録された知識、情報の発生・伝達・処理・蓄積・利用およびこれらのための組織を扱う学問分野に対して与えられた名称の変遷とそれらの内包については次の論文を見よ。
Wellisch, Hans. "From information science to informatics: A terminological investigation." *Journal of librarianship*, vol. 4, July 1972, p. 157-87.
Brookes, B. C. "From Information science to informatics (Correspondence)," *Journal of librarianship*, vol. 4, Oct. 1972, p. 272-4.
 - 5) Brookes, B. C. Information science (excluding IR). <Whatley, *op. cit.*> p. 138.
 - 6) *Ibid.*, p. 146.
 - 7) Brookes, B. C. "Numerical methods of bibliographic analysis," *Library trends*, vol. 22, July 1973, p. 18.
 - 8) システム分析には広狭二つの意味がある。広義には、これは、ある機能を果たしているシステムの効率を改善する目的を持って（あるいは、ある定められた機能を果たすべきシステムをつくり上げる目的を持って）、その（あるいは、類似の）システムを研究し、その結果に基づいて必要な改善策を講じて、予め設定した機能を円滑に果たするようなシステムをつくりあげること、と定義できる。この場合、システム分析の過程は一般に次のステップから成り立っている：(1) 目標設定、(2) データ収集、(3) 分析、(4) システム設計、(5) インプリメンテーション、(6) 検査・フィードバック。
狭義には、これは、現在稼動しているシステムの吟味のみをさし、この場合には、上の(2)、(3)の二つのステップのみで作業は終る。しかし、システム分析の具体的手順を述べる場合は別として、一般にシステム分析という場合は広い意味で使われるのが普通であり（これについては下記の文献を見よ）、ここでも一般の用法に従っている。
Robinson, F., *et al. Systems analysis in libraries*. (Symplegades, no. 2). Newcastle upon Tyne, Oriel Press. p. 1-2.
Chandor, Anthony, *et al.* コンピュータ用語辞典. (坂井利之監訳) 東京、講談社、1972. p. 426-31. (ブルーバックス B-195).
 - 9) この意味で、“情報検索には一般論はなく、あるのはただ特論だけだ、というのが、前々からの筆者の持論です”という細井の意見に、筆者は賛成する。ここで彼の言う情報検索とは、いわゆる、情報検索システムのことである。
細井 勉. “論理的にみた定理検索の可能性,” *数理科学*, no. 129 (特集：情報検索), Mar. 1974, p. 66.
 - 10) 前川は、「今後の図書館統計研究の方向と問題」として、次の点をあげている。（ただし、原文のままではない）
1. 評価と統計の視点を定める。このためには図書館の本質的機能を各館種ごとに明確にする必要がある。
2. 統計以前の問題として、数の数え方が全国の図書館で一定ではない。数え方の実態調査と統一が急務である。
3. 現在は、さまざまな率、割合などを出して、その意味、相互関係などを研究している段階であるが、さらに進めて、図書館運営に数学的な考え方をとり入れるべきである。
前川恒雄. 図書館の管理. <武田虎之助編, 図書館学学習の手引. 東京、日本図書館協会、1973> p. 144.
 - 11) Morse, Philip M. *Library effectiveness: A systems approach*. Cambridge, Mass., M. I. T. Press, 1968. p. 84.
 - 12) Buckland, M. K. and Woodburn, I. *Some implications for library management of scattering and obsolescence*. (University of Lancaster Library occasional papers, no. 1) Lancaster, University of Lancaster Library, 1968.
Buckland, M. K., *et al. Systems analysis of a university library: Final report on a research project*. (University of Lancaster Library occasional papers, no. 4) Lancaster, University of Lancaster Library, 1970.
 - 13) 前項で述べたシステム分析の手順に従えば、図書館・情報システムを設計・運営する際に、利用者研究は、(1) システムの目標を設定するための利用者の情報要求および利用パターンの識別、(2) 目標の変更、システムの機能の低下などによるシステムの目標と達成のずれを発見して、それをシステムにフィードバックするための、利用者のニーズ・利用の変化および満足度の識別、の手段として欠かすことのできないステップであるから、利用者研究の流行はシステム分析の発展の中を含めて考えることもできよう。しかし、現在多くの

- 国々で行なわれている利用者研究が、すべて具体的なシステムの設計を念頭においているとはいえず、むしろ、科学者の一般的特性を知るための独立した研究テーマと考えた方が良いと思われる。
- 14) Brittain, J. M. *Information and its users: A review with special reference to the social sciences*. Bath, Bath University Press, 1970. p. 125.
- 15) 例えば、スレイターは、現在英国で行なわれている利用者研究全般を次のように批判している：“この（利用者・図書館調査の）パターンは現在まだ対称的でもないしバランスのとれたものでもない。……研究の方向についていかなる組織化された同意もないことも研究努力のこの点在的な分布の原因となっている。マスタープランの欠如はこれまでに行なわれた多くの研究の結果の間の矛盾にも反映されている。孤立的な諸問題あるいは特定の諸グループがいくつかの異なる時期に異なる環境においてそしてしばしば不必要にも個別的な方法で研究されてきた。その結果生み出された文献から何らかの一般的結論を引き出すことはそれ自体一つの技術となった。”
- Slater, Margaret. User and library surveys. <Whatley, *op. cit.*> p. 232.
- 16) Cole, F. J. and Eales, Nellie B. “The history of comparative anatomy. Part I.—A statistical analysis of the literature.” *Science progress*, vol. 11, Apr. 1917, p. 578.
- 17) *Ibid.*, p. 579.
- 18) 彼の研究成果はケンブリッジ大学における講演で発表され、後に単行書として出版された。
- Hulme, E. Wyndham. *Statistical bibliography in relation to the growth of modern civilization: Two lectures delivered in the University of Cambridge in May 1922*. London, Grafton, 1923.
- ただし、原著が入手できなかったので、彼の研究内容の紹介は次にあげる上掲書の書評および2年後に発表された彼の短い論文によった。
- “A Survey of scientific literature,” *Nature*, vol. 112, Oct. 20, 1923, p. 585-6.
- E. W. H. “Output of scientific papers,” *Nature*, vol. 116, July 25, 1925, p. 129-30.
- （この著者がヒュームであることは上の著書への言及によって明らかである。）
- 19) 彼はまず *ICSL* の1ページあたりの平均記載文献数から科学文献数の年次変化および国別の生産数を推定して、科学文献は、周期的な変動をくり返しながらも、全体として急激に増加していること、またこれらは少数の国に偏っていることなどを社会や産業の発展の歴史と関連づけて説明した。
- また、1561年以來のイギリスの特許に関する統計から、特許明細書の導入が与えた影響、とくに特許の実用化からそれに関する文献の発表までの時間的遅れ、その結果としての産業界における文献利用の少なさを明らかにした。
- 20) もっともまったく無視されていたわけではない。次の著書を見よ。
- Malclès, Louise-Noëlle. *Bibliography*. Translated by Theodore Christian Hines. New York, Scarecrow, 1961. p. 8.
- 21) Lotka, Alfred J. “The frequency distribution of scientific productivity,” *Journal of the Washington Academy of Sciences*, vol. 16, June 19, 1926, p. 317-23.
- なお、ロトカの法則については次の著書も参照されたい。
- Price, Derek J. de Solla. *Little science, big science*. New York, Columbia University Press, 1963. Chap. 2.
- 22) Lotka, *op. cit.*, p. 323.
- 23) 人文科学の分野においても、研究者の生産性に関して、ロトカの法則が成立することが確かめられている。
- Murphy, Larry J. “Lotka’s Law in the humanities?” *Journal of the American Society for Information Science*, vol. 24, Nov.-Dec. 1973, p. 461-2.
- ただし、この論文の調査対象は技術史の専門雑誌である *Technology and Culture* なので厳密に人文科学と言えるかどうか疑問である。
- 24) 例えば、次のような論文・著書がある。
- Stevens, Rolland E. *Characteristics of subject literatures*. Chicago, Association of College and Reference Libraries, 1953. 12 p. (ACRL monograph, no. 6).
- Brown, Charles Harvey. *Scientific serials*. Chicago, Association of College and Reference Libraries, 1956. 189 p. (ACRL monograph, no. 16).
- Törnudd, Elin. Study on the use of scientific literature and reference services by Scandinavian scientists and engineers engaged in research and development. <*Proceedings of the International Conference on Scientific Information, Washington, D. C., Nov. 16-21, 1958*. Washington, D. C., National Academy of Science—National Research Council. 1959.> Vol. 1. p. 19-75.
- 25) ブラッドフォードの法則は、本来、引用文献の分散について定められた法則ではない。これについてはIVを見よ。
- 26) Gross, P. L. K. and Gross, E. M. “College libraries and chemical education,” *Science*, vol. 66, Oct. 28, 1927, p. 385-9.
- 27) *Ibid.*, p. 389.

- 28) *Ibid.*, p. 387.
- 29) Fussler, Herman H. "Characteristics of the research literature used by chemists and physicists in the United States," *Library quarterly*, vol. 19, Jan. 1949, p. 19-35. (Part I); *Library quarterly*, vol. 19, Apr. 1949, p. 119-43 (Part II).
- 30) *Ibid.*, p. 140.
- 31) *Ibid.*, p. 141.
- 32) Gross and Gross, *op. cit.*
- 33) 物理学, 化学における言語の分布の詳細な調査の結果は注 29) で引用したフスラーの論文に見られる。また, 注 24) で引用したスティーブンスの論文には多くの調査結果が表にまとめられている。
- 34) Voigt, Melvin J. "Scientific periodicals as a basic requirement for engineering and agricultural research," *College and research libraries*, vol. 8, part II, July 1947, p. 354-9, 375.
- 35) Fussler, *op. cit.*
- 36) *Ibid.*, p. 122.
- 37) この表はフスラーの上記論文中的数据をもとに作成された。調査対象は 1939 年に刊行された雑誌である。
- 38) なお, これとは逆方向の研究, すなわち, あるアイデアまたは理論が各分野に伝播していくのにどの位の時間を要したかについての研究が, 他分野の出版物によるシャノンの通信理論の引用を例にとって, ダーリングによって行なわれている。彼によると, シャノンの論文は発表から 7 年後までに 9 つの領域の出版物に引用されたそうである。つぎの論文を見よ。
Crane, Diana. "Information needs and uses," *Annual review of information science and technology*, vol. 6, 1971, p. 16.
- 39) Gross and Gross, *op. cit.*
- 40) 例えば, グロスらは, アメリカにおける地質学の文献についての調査結果から, アメリカの地質学者にとっては「米国およびカナダの政府報告書が第一に重要である」との結論を得ている。
Gross, P. L. K. and Woodford, A. O. "Serial literature used by American geologists," *Science*, vol. 73, June 19, 1931, p. 660-5.
- 41) Sengupta, I. N. "Recent growth of the literature of biochemistry and changes in ranking of periodicals," *Journal of documentation*, vol. 29, June 1973, p. 202.
当然のことながら, これは政府機関の補助金を受けて行なわれる研究がふえ, これらの成果がモノグラフの形で出版されたことにもよる。
- 42) 大きな分散は境界領域においてとくに顕著に現われることは容易に推測できる。Sengupta, *op. cit.*, p. 197.
- 43) ブラッドフォードが彼の得た「法則」を述べた論文を発表したのは 1934 年である。しかし, この論文は, 主として実験の方法と結果を述べたものである。彼が実験を行なうにいたった動機・背景の詳細をこの論文から知ることはできない。このため, 本節および次節の記述は, 主として, 彼の著書 *Documentation* によっている。
Bradford, S. C. "Sources of information on specific subject," *Engineering*, vol. 137, Jan. 26, 1934, p. 85-6.
Bradford, S. C. *Documentation*. 2d ed. London, Crosby Lockwood, 1953. Chapters IX and X.
- 44) このような「書誌的混沌」から脱け出すために, すなわち彼の言う「完全なドキュメンテーション」(complete documentation)を実現するために, ブラッドフォードが考えたのは書誌サービスの統合であった。彼は, そのための前提として, 国際十進分類法 (UDC) によってすべての論文を分類すべきであると主張している。この主張は彼の論文にくり返し現われ, この点だけを見ると, 彼も当時の国際ドキュメンテーション協会の代弁者にすぎなかったとの評価も生まれてくるかもしれない。さらに, 彼が国際ドキュメンテーション協会の英国支部とも言うべき「英国国際書誌学会」(British Society for International Bibliography)の有力なメンバーであったことはこの評価のうらづけとなろう。しかし, 彼の活動が UDC の普及のみでなかったことは, 科学博物館図書館において彼の指導の下に行なわれた「補助貸出し計画」(Supplementary Loan Scheme), 彼の国立科学図書館設置提案, などを見てもわかることである。これについての詳しい議論は本論の主題から外れるので省略する。
- 45) SML は科学博物館 (Science Museum) の 9 つの department の一つであるので, SML は科学博物館図書館部, 彼の肩書は部長の方が適当かもしれない。しかしここでは, BML の日本での呼称に準じておく。図書館長 (Keeper) は博物館長 (Director) の監督を受ける。
- 46) Bradford *Documentation*, *op. cit.*, p. 146.
- 47) *Ibid.*, p. 147-8.
- 48) 上にあげた論文以外の論文をも含めて, 彼の論文にはグロスらの論文に言及した箇所はないので, 彼はグロスらの研究を知らなかったと思われる。しかし, 仮に彼がグロスらの研究を知っていたとしても, 前項で述べたように, ある主題の論文がどのように「他の」雑誌に分散しているかを知らうとした彼の研究目的にグロスらの方法が役に立ったとは思われない。
- 49) この書誌は, 科学博物館図書館で受け入れている抄録誌を調査し, それ以前から集められ UDC 標数をつけて分類されていた論文をもとに SML で

作成された文献目録を調べ、毎日受け入れる「他の」雑誌を精査して作成されたもので、完全とは言えないまでも、かなりよく作られた主題書誌であったと言える。

- 50) Bradford, "Sources of information on specific subject," p. 86.あるいは、Bradford, *Documentation, op. cit.*, p. 150.
- 51) ブラッドフォードは、Dを応用地球物理学については2倍、潤滑については5倍して、いずれについてもほぼ同じ傾きの曲線を描いているが、これはこのグラフの意味を誤解させるおそれがあるので、ここでは縦軸はDの値をそのままとっている。なお、ブルックスはこのようなグラフをbibliographと呼んでいる。
Brookes, B. C. "The complete Bradford-Zipf 'bibliograph'," *Journal of documentation*, vol. 25, Mar. 1969, p. 58-60.
- 52) 「ブラッドフォードの分散則」という名称はウィッカーリーによって与えられた。
Vickery, B. C. "Bradford's Law of Scattering," *Journal of documentation*, vol. 4, Sept. 1948, p. 198-203.
- 53) Bradford, *Documentation, op. cit.*, p. 154.
- 54) *Ibid.*
- 55) *Ibid.*, p. 148.
- 56) Garfield, Eugene. "The mystery of the transposed journal lists—Wherein Bradford's Law of Scattering is generalized according to Garfield's Law of Concentration (Current comments)," *Current contents*®, vol. 3, Aug. 4, 1971, p. 5-6.
- 57) 中村千里. "試験場の図書館 (2)," *農業技術*, vol. 28, 1973, p. 282.
- 58) Garfield, *op. cit.*, p. 5.
- 59) *Ibid.*, p. 6.
- 60) Sengupta, *op. cit.*, p. 211.
- 61) Vickery, *op. cit.*
- 62) Brookes, B. C. "The derivation and application of the Bradford-Zipf distribution," *Journal of documentation*, vol. 24, Dec. 1968, p. 257.
なお、これは直線部分についてのみ成り立つ式であり、核の部分は
$$R(n) = an^b$$
で近似できることをブルックスは示しているが(注51)にあげた論文, p. 58), 核を一つの単位と考えれば、その「微細構造」を知る必要はない。
- 63) この式の導出については、次の論文を見よ。
Leimkuhler, Ferdinand F. "The Bradford distribution," *Journal of documentation*, vol. 23, Sept. 1967, p. 197-207.
- 64) つぎの論文も見よ。
Wilkinson, E. A. "The ambiguity of Brad-

ford's Law," *Journal of documentation*, vol. 28, June 1972, p. 122-130. および "Errata" *ibid.*, vol. 28, Sept. 1972, p. 232.

- 65) Vickery, *op. cit.*, p. 202.
- 66) Wilkinson, *op. cit.*, p. 126.
- 67) Cole, P. F. "A new look at reference scattering," *Journal of documentation*, vol. 18, June 1962, p. 58-64.
- 68) Goffman, William and Warren, Kenneth S. "Dispersion of papers among journals based on a mathematical analysis of two diverse medical literatures," *Nature*, vol. 221, Mar. 1969, p. 1205-7.
- 69) Kendall, M. G. "The bibliography of operational research," *Operational research quarterly*, vol. 11, Mar./June 1960, p. 31-6.
- 70) 計数作業の詳しい手順については、例えば、次の論文を見よ。中村千里. "農学分野における文献情報利用調査の展望と Citation Counting の展開," *Library and information science*, no. 8, 1970, p. 56-7; Sengupta, *op. cit.*, p. 195-7.
- 71) Martyn, John and Gilchrist, Alan. *An evaluation of British scientific journals*. London, Aslib, 1968. (Aslib occasional publication, no. 1).
- 72) Sengupta, *op. cit.*
- 73) Sengupta, I. N. "Impact of scientific serials on the advancement of medical knowledge: An objective method of analysis," *International library review*, vol. 4, Apr. 1972, p. 169-95.
- 74) Bradford, "Sources of information on specific subject," *op. cit.*
- 75) Kendall, *op. cit.*
Goffman and Warren, *op. cit.*
- 76) Martyn and Gilchrist, *op. cit.*, p. 1.
- 77) Brittain, J. Michael and Line, Maurice B. "Sources of citations and references for analysis purposes," *Journal of documentation*, vol. 29, Mar. 1973, p. 72-80.
- 78) 引用文献の共通性を利用した文献の分類はケスラーによって考えられた。これについてはVでふれるが、雑誌は論文の集合であるから、文献の分散による制限はあるが、引用文献の共通性によって雑誌を分類することもまたある程度可能である。
- 79) 著者の自己引用を知ることができるという特長はそのままこれが含まれていることによって利用パターンがゆがめられているという欠点にもなる。
- 80) 引用文献が実際の利用を反映していないと言われる理由はいくつかある。第一に、引用文献にはその論文を書くために読まれた文献だけではなく、長い間蓄積されてきた文献に関する知識が反映している。社会科学の引用文献(雑誌のみ)の半減期が6年であるのに対して、NLLに貸出要求の

あった社会科学雑誌の半減期は3.5年であった、というアールらの調査結果はこの事実を裏書きしている。

Earle, Penelope and Vickery, B. C. "Social science literature use in the UK as indicated by citations," *Journal of documentation*, vol. 25, June 1969, p. 123-41.

第二に、余りにもよく引用される有名な論文であったり、内容はおぼえているが掲載誌・巻・号などがあいまいであったために、読まれたのに引用されない論文がある。(論文が引用されない理由については次の論文を見よ。)

Garfield, Eugene. "Uncitedness III — The importance of *Not* being cited (Current comments)," *Current contents*®, vol. 5, Feb. 21, 1973, p. 5-6.

第三に、著者が読んでいない論文が参考文献としてあげられることもある。レビュー誌、抄録誌の発達、図書館・情報サービスの充実は、これをますます容易にする。

- 81) 下記の報告はこの事実を裏書きしている。著者は、農学関係雑誌の引用文献調査において、『青森りんご試報告』に外国雑誌からの引用が多いのは、この試験場所で蔵する雑誌のタイトル数が多いこと、外国雑誌の比率が高いこと、さらに、図書館活動が他の同種機関より活発であること、などによるのであろう、としている。

中村千里, 北村晴夫. "農学部門における文献情報の計量," 図書館管理運営に関する研究会報告(昭和42年度), 農林水産技術会議事務局調査資料課, 1968. 3, p. 20.

- 82) 調査対象としてどのような雑誌を選ぶかは難しい問題である。さらに、ある雑誌に掲載される論文は必ずしもすべてがある定められた主題を扱っているとはいえないことは分散則の系として当然出てくるので、これをすべて関連論文と考えるのは自己矛盾であるとも言える。

- 83) Brodman, Estelle. "Choosing physiology journals," *Bulletin of the Medical Library Association*, vol. 32, Oct. 1944, p. 479-83.

- 84) この計算は次のスピアマンの公式による。

$$\rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

ここで、

ρ : 順位相関係数

D: 二つの異なる基準によって定められた順位の差

N: 順位づけされたペアの数

- 85) *Ibid.*, p. 482.

- 86) ヴィッカリーは、(1) サンプルとして選んだ世界の科学技術雑誌の中の論文の著者による引用(SCIを利用)、(2) サンプルとして選んだ英国の著者による引用、(3) NLLに出された貸出し要

求(実際に貸出された回数ではない)、(4) 各々の雑誌を所蔵している英国の図書館の数、(5) 一年間に掲載された論文とレターの数、などの5つの「指標」(indicator)によって上位89誌を求めている。これら5つのすべてに共通なのはわずか12誌だったそうである。つぎの論文を見よ。

Vickery, B. C. "Indicators of the use of periodicals," *Journal of librarianship*, vol. 1, July 1969, p. 170-82.

なお、引用と貸出しの二つの方法による結果の比較は、注80)にあげたアールらの論文を見よ。

- 87) Cole, P. F. "Journal usage versus age of journal," *Journal of documentation*, vol. 19, Mar. 1963, p. 1-11. 彼は「 x 年より古い参考文献の数 R_x 」と「参考文献の総数 RT 」との関係を示したが、式の形は同じになる。

- 88) Burton, R. E. and Kebler, R. W. "The 'half-life' of some scientific and technical literature," *American documentation*, vol. 11, Jan. 1960, p. 18-22.

ところで、パートンらがこの用語を提案したときに、文献の寿命に関して得た経験式は式(10)のような形ではなく、次のような形であった。

$$y = 1 - (ae^{-x} + be^{-2x})$$

ここで、

y は論文の(若い方からの)累積和の全論文に対する比、

x は論文の年令、

を表わし、 $x=0$ のとき $y=0$ より $a+b=1$

となる。

この式は x が小さい部分で見られるずれを外挿によって補正したものであり、それゆえ、出版から利用までの時間的遅れは考慮に入っていない。

彼らは、この式が二つの成分から構成されているのは、大部分の分野には古典的な文献と短命な(ephemeral)な文献の二つの集合がありうることによると説明している。もしこの二成分を分離できれば興味ある結果が得られると思われるが、これを検証することはデータのバラツキのため困難であり、彼らの解釈は単なる推測にとどまらざるを得ない。

- 89) Bourne, Charles P. Some user requirement stated quantitatively in terms of the 90 percent library. <Kent, A. and Taulbee, O. E., eds. *Electronic information handling*. Washington, D. C., Spartan Books, 1965> p. 107.

- 90) ワイスは生物学関係の二つの雑誌、生理学および動物学・昆虫学の二つの分野の主要な雑誌の引用文献を調べて、その年令分布曲線を得、これより、知識も、高等動物の器官と同じく、以前の知識を利用して成長していき、古い知識は新しい知識に同化されたところで死滅すると思った。Weiss, Paul. "Knowledge: A growth process," *Sci-*

- ence, vol. 131, June 10, 1960, p. 1716-9.
- 91) Goffard, S. James and Windle, Charles D. "Life of scientific publications (Letter)," *Science*, vol. 132, Sept. 2, 1960, p. 625.
- 92) 文献の成長および寿命が指数関数で近似でき、それぞれのパラメータを g , λ とすると、ある年度に発表された論文全体について、1論文あたりの利用度は t とともに

$$y = y_0 e^{-(\lambda - g)t}$$
に従って変化する。ここで $\lambda = g$ であれば y は定数となり、論文の相対的価値は年令によって変わらないことになる。
- 93) Price, Derek J. de Solla. "Networks of scientific papers," *Science*, vol. 149, July 30, 1965, p. 510-5.
この論文の邦訳は次に付録として収められている。Price, Derek J. de Solla, リトル・サイエンス, ビッグ・サイエンス, 島尾永康訳, 大阪, 創元社, 1970. なお彼は最近の論文がより多く利用されることを「即時性の因子」(Immediacy Factor)と名づけている。
- 94) Line, M. B. "The 'half-life' of periodical literature: Apparent and real obsolescence," *Journal of documentation*, vol. 26, Mar. 1970, p. 46-52. 次にあげたこの論文に対するヴィッカリーの注記を見よ。Vickery, B. C. *Journal of documentation*, vol. 26, Mar. 1970, p. 53-4, さらに, Line, M. B. "The half-life of periodical literature (Letter to the editor)," *Journal of documentation*, vol. 26, Sept. 1970, p. 266.
- 95) ところが、ここで彼が用いた成長は、雑誌数の増加ではなく、論文数の増加である。彼は、"この二つの成長率はもちろん同じ値あるいはきわめて近い値になるであろう"と述べているが、1雑誌あたりの収録論文数が増加していることは良く知られた事実であり、この二つの成長率は明らかに異なる値をもつ。なお、注 94) であげたレターでは、「修正された半減期」と「真の半減期」との区別は非常にあいまいになってしまっている。
- 96) Vickery, *Journal of documentation*, vol. 26, 1970, p. 54.
- 97) Brookes, "Numerical methods of bibliographic analysis," p. 18.
- 98) Salton, Gerard. "Automatic indexing using bibliographic citations," *Journal of documentation*, vol. 27, June 1971, p. 98-110.
- 99) Kessler, M. M. "Comparison of the results of bibliographic coupling and analytic subject indexing," *American documentation*, vol. 16, July 1965, p. 223.
- 100) *Loc. cit.*
- 101) Kessler, M. M. "Bibliographic coupling between scientific papers," *American documentation*, vol. 14, Jan. 1963, p. 11.
- 102) Small, Henry. "Co-citation in the scientific literature: A new measure of the relationship between two documents," *Journal of the American Society for Information Science*, vol. 24, July-August 1973, p. 265-9.
- 103) Price, Derek J. de Solla. Citation measures of hard science, soft science, technology, and nonscience. (Nelson, Carnot E. and Pollack, Donald K., eds. *Communication among scientists and engineers*, Lexington, Mass., Heath Lexington, 1970), p. 3-22.
- 104) Earle and Vickery, *op. cit.*
- 105) Goffman, W. and Newill, Vaun A. "Generalization of epidemic theory: An application to the transmission of ideas," *Nature*, vol. 204, Oct. 17, 1964, p. 225-8.
- 106) Goffman, W. and Harmon, G. "Mathematical approach to the prediction of scientific discovery," *Nature*, vol. 229, Jan. 8, 1971, p. 103-4.
- 107) Goffman, W. "Mathematical approach to the spread of scientific ideas—The history of mast cell research," *Nature*, vol. 212, Oct. 29, 1966, p. 449-52.
- 108) Donohue, Joseph C. *Understanding scientific literatures: A bibliometric approach*. Cambridge, Mass., MIT Press, 1973.
- 109) 山田圭一, 大滝栄一. "科学研究のライフ・サイクル—リサーチ・オン・リサーチへの一つのアプローチ," *自然*, no. 309, 1971, 12, p. 68-76.
- 110) Goffman and Harmon, *op. cit.*
- 111) 彼らはこの「予測」を次のような方法で行なった。まずこの方法は次のような仮定に基づいている。
(1) ある学問分野の、時刻 $t + \Delta t$ における、発展速度は時刻 t におけるその分野の刺激行動の相対的な強さと、他の諸分野との比較におけるその刺激行動の力点の、時間 Δt の間の、変化との函数であるという意味で、科学発展のプロセスはグローバルな刺激物 (stimulators) の刺激行動の結果であると考えられる。
(2) このような刺激物として、出版物、学会、会議、個人的接触などがあるが、これらに関するデータのうちの最も入手しやすいのは出版物中の引用である。
(3) 原論文、引用論文はそれぞれどの分野に属するかを定めることができるから、すべての論文についてこのような分類を行なえば、ある分野の論文の引用文献はどのような分野の論文から構成されているかがわかる。
(4) 以上より、引用論文および原論文の諸分野

への分散は、それぞれ時刻 $t, t+dt$ における、各分野の繁栄の度合いを示しており、分野 j に属する論文で分野 i に属する論文を引用している論文が P_{ij} 篇あるとき、この

値を規格化した $p_{ij}(=P_{ij}/\sum_{j=1}^n P_{ij})$ は分野間

の遷移確率と考えることができる。

これらの仮定より、初期の科学研究の繁栄度を引用論文の分野別の割合を元とするベクトル $\mathbf{p}(0)$ で表わし、上の p_{ij} を元とする行列 P を遷移行列とした。「予測」は P の極限分布と $\mathbf{p}(0)$ とを比較することによって行なっている。

Zunde, P. and Slamecka, V. "Predictive models of scientific progress," *Information storage and retrieval*, vol. 7, 1971, p. 103-9.

- 112) このことを具体例をあげて証明してみよう。簡単のために、分野を二つとし（社会科学と自然科学としてもよい）、これを 1, 2 で表わす。

p_{ij} を i の論文が j の論文に引用される確率とすると、遷移確率 P は、

$$P=(p_{ij})=\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

この極限分布 $\mathbf{p}(\infty)$ は、 $p_{i1}+p_{i2}=1$ であるから、

$$\mathbf{p}(\infty)=\left(\frac{1-p_{22}}{2-(p_{11}+p_{22})}, \frac{1-p_{11}}{2-(p_{11}+p_{22})}\right)$$

となり、結局、彼らの方法に従えば、他分野への貢献度の大きい分野ほど、将来衰えていくという奇妙な結論に到達してしまう。

- 113) 証明は次の通りである。

最下位の雑誌は、一般に、関連論文を 1 篇しか含んでいないと考えられるので、

$$R(N)-R(N-1)=1$$

となる。一方、(11) 式より、

$$\begin{aligned} R(N)-R(N-1) \\ =k \ln \frac{N}{s}-k \ln \frac{N-1}{s} \end{aligned}$$

$$=-k \ln \left(1-\frac{1}{N}\right)$$

であるが、

$$\ln(1-x)=-\sum_{n=1}^{\infty} \frac{x^n}{n}$$

および $N \gg 1$ より、

$$R(N)-R(N-1) \approx \frac{k}{N}$$

故に、

$$k \approx N$$

Brookes, B. C. "Bradford's Law and the bibliography of science," *Nature*, vol. 224, Dec. 6, 1969, p. 953-6.

- 114) ここで、個々の雑誌の寿命曲線も指数型になることを仮定していることに注意したい。引用文献の寿命曲線が指数関数曲線になることはこれまで多くの実験により確認されているが、これは文献全体の総和により得られたパターンであって、このことが個々の雑誌についてもいえるとは即断できない。しかし、この仮定が不当でないことは、一つにはグロスらの調査結果などにより、一つには廃棄の問題を考える場合に、まず第一に重要なのは平均的なふるまいを知ることであることにより、認められるであろう。

- 115) なお、中村の用いた「情報率」はこの p_1 と引用文献の総数に対する雑誌論文の割合との積である。

中村, *op. cit.*

- 116) これは、ブルックスが彼のいう「最適 $P\%$ 図書館」を考える際においた仮定と同じである。

Brookes, B. C. "Optimum $P\%$ library of scientific periodicals," *Nature*, vol. 232, Aug. 13, 1971, p. 458-61.

- 117) Brookes, "The complete Bradford-Zipf 'bibliograph'," *op. cit.*

- 118) Donohue, *op. cit.*