

我が国の各学問分野における論文引用の傾向分析

○山下泰弘*, 西澤正己+, 孫媛+, 根岸正光+

Tendency analysis of article citations in Japanese academic disciplines

○Yasuhiro YAMASHITA*, Masaki NISHIZAWA+,
Yuan SUN+, Masamitsu NEGISHI+

Abstract

In recent years, research evaluation is becoming one of the major topics of governmental science policy. Along with this trend, citation analyses attract great attention as a tool for quantitative measure of quality of research output.

We report some results of an analysis on the characteristics of citations in academic disciplines, based on the ISI citation statistic database: National Citation Report (NCR). This database contains bibliographies of articles written by authors affiliating to Japanese institutes from January 1981 to June 1997, with yearly citation counts.

First, we analyze frequency distributions of citations by academic disciplines, and then investigate changes of citation counts per article through the years after the publication. Finally, we classify all disciplines into four types by the two indices: number of papers and cited times per article.

1. はじめに

近年、大学等研究機関において研究評価の必要性が認識されつつあり、多くの機関において自己点検・評価活動が実施されている。研究活動の評価における重要な定量的指標としては、研究論文数や被引用数などが従来からよく用いられてきた。しかしながら、それぞれ分野ごとに発行しやすさや、引用されやすさの差異が大きく、またその状況が十分に把握されていない。今後、多様な分野の研究者の集まる総合的学部・研究科や学際的プロジェクト等についての評価を実施する上で、これら定量的指標について、より分野横断的な視点からの情報が必要になるものと考えられる。

本研究では、このような観点から研究評価における基礎情報である引用状況について、我が国の学界全体を網羅的に扱った実証分析を行い、今後の評価活動に有用な基礎データを作成する。

2. 使用するデータ—National Citation Report (for Japan)—

本研究では、データソースとして、ISI社の引用統計データベース“National Citation Report (for Japan)” (NCR) を使用する。NCRは、ISIにより編集・提供されているSCI (Science Citation Index、自然科学分野) およびSSCI (Social Science Citation Index、社会科学分野)、A&HCI (Art & Humanities Citation Index、人文科学分野) の3つの引

用統計データベースから、特定国の論文のみを抽出して作成されている。本研究で使用する NCR は、1981 年 1 月から 1997 年 6 月までの 16 年半の間に ISI 社の引用統計データベースに収録された 853,323 件の「日本の論文」が収録されており、被引用数についても 1997 年 6 月まで計測されている。本研究では、発行直後の論文の被引用数は一般に少ないことを考慮し、以下では 1994 年までに発行された論文 664,190 件を対象として分析を行なう。また引用数の経年変化についての分析については、発行後の被引用度の経年変化を見るため、さらにマージンをとって 1981 年から 1990 年までの 10 年間に発行された論文を対象とする。

分野については、原データベースでは約 100 分類の category code が付与されているが、分析に使用するには細かすぎるので、National Science Indicator (Standard 版) で使用されている 26 分野に再分類した（無区分を含む）。さらに、個別分野の分析に際しては、それらを数物系、生物・化学系、人文・社会科学系の 3 分類に統合した。なお、本研究で用いている分野分類の詳細については別の報告（孫 2000, p.14）を参照されたい。

3. 分析結果

3. 1 学問分野別の論文数と被引用数

各分野の 94 年までに発行された論文 1 件当り被引用回数（以下被引用度と呼ぶ）を表 1 に示す。全体では、論文数は 664,190 件で 1 件当り 8.2 回引用されている。論文数は、Clinical Medicine (121,176 件)、Chemistry (115,611 件)、Physics (95,045 件)、の順に多い。一方、人文・社会科学系諸分野はいずれも論文数が少ない。特に Law は 79 件、Education は 117 件に過ぎず、統計分析が困難なため、以下では全体についての分析には含めるが、個別分野についての分析においては考察対象からはずしてある。

表 1 我が国の分野別論文数と引用度（1981-1994）

分野		論文数	被引用度	分野		論文数	被引用度
All		664,190	8.2	生命科学系	Microbiology	16,991	12.5
数学・物理系	Astrophysics	3,805	12.8		Molecular Biology & Genetics	19,891	14.4
	Chemistry	115,611	9.1		Neurosciences	22,780	11.3
	Computer Science	5,390	2.6		Pharmacology	27,647	6.7
	Engineering	54,732	4.4		Plant & Animal Sciences	34,385	5.2
	Geoscience	7,257	7.7		Psychology/Psychiatry	3,940	3.2
	Materials Science	34,106	4.4	人文・社会科学系	Economics & Business	1,606	2.9
	Mathematics	6,147	3.7		Education	117	1.8
	Physics	95,045	9.0		Law	79	1.9
生命科学	Agricultural Sciences	22,936	5.5		Social Sciences, General	3,391	1.7
	Biology & Biochemistry	74,152	14.6		Art & Humanities	2,515	0.2
	Clinical Medicine	121,176	7.0		Multidisciplinary	8,738	3.4
	Ecology / Environment	5,723	4.8		Not Specified	26,798	2.2
	Immunology	8,526	16.8				

94年までに発行された論文の被引用度は、全体では8.2となっている。分野別に見ると、被引用度が高いのは Immunology (16.8)、Biology & Biochemistry (14.6)、Molecular Biology & Genetics (14.4)、Astrophysics (12.8)、Microbiology (12.5)、Neurosciences (11.3) などであり、Astrophysics 以外はいずれも生命科学系分野となっている。一方、人文・社会科学系分野は概して被引用度が低く、Art & Humanities が 0.2、Social Science, General が 1.7 と生命科学系諸分野と比べてかなりの差が見られる。被引用回数は、原データベースに収録されている雑誌についてのものであり、対象分野の収録雑誌数に比例して多くなりやすいため、一概にこれらの分野の被引用回数が多いとは言い切れないが、論文数の多い分野についてはある程度分野の引用傾向を反映しているものと考えられる。

3. 2 被引用度数分布

94年までに発行された論文の被引用度数分布を見ると、全体では被引用回数0の論文が一番多く、26.5%に及んでいる。以下1回が8.7%、2回が7.0%、3回が5.7%と漸減している。

個別分野については、ここでは便宜的に被引用回数0の論文の比率（相対度数）が20%以上の分野を多死型、20%未満の分野を少死型と類別して考える。

数学・物理系分野では、少死型分野には Astrophysics、Geoscience、Physics の3分野が含まれる。これらの分野の論文は、その80%以上が回数の多少はあっても何らかの形で引用されていることになる。一方数学・物理系の多死型分野には、Computer Science、Engineering、Materials Science、Mathematics の4分野が含まれる。これらの分野においては、発行された論文の20%以上が、まったく引用されていない。特に、Computer Science においては、発行された論文の過半数（54%）が引用されていない。

生命科学分野では、Biology & Biochemistry、Chemistry、Immunology、Microbiology の4分野が少死型に含まれる。これらは Chemistry を除くと先ほど挙げた被引用度の高い分野に含まれ、全体的に引用されやすい分野と見ることができる。多死型分野には、Agricultural Sciences、Clinical Medicine、Ecology / Environment、Molecular Biology & Genetics、Neurosciences、Pharmacology、Plant & Animal Sciences、Psychology/Psychiatry の8分野が含まれる。これらのほとんどは比較的被引用度の低い分野であるが、Molecular Biology & Genetics、Neurosciences は総合的には被引用度の非常に高い分野である。したがってこれらの分野は一部の高被引用論文により全体の平均が引き上げられていると考えられる。また、これらの分野の中で Psychology & Psychiatry は極めて被引用回数0の論文が多くなっており、どちらかという後述の人文・社会科学分野に近い性質を持っているものと考えられる。

人文・社会科学分野では、少死型分野が存在せず、いずれも多死型分野となっている。特に、Art & Humanities は約90%の論文が被引用回数0となっており、論文が引用されるということは極めてまれな現象である。

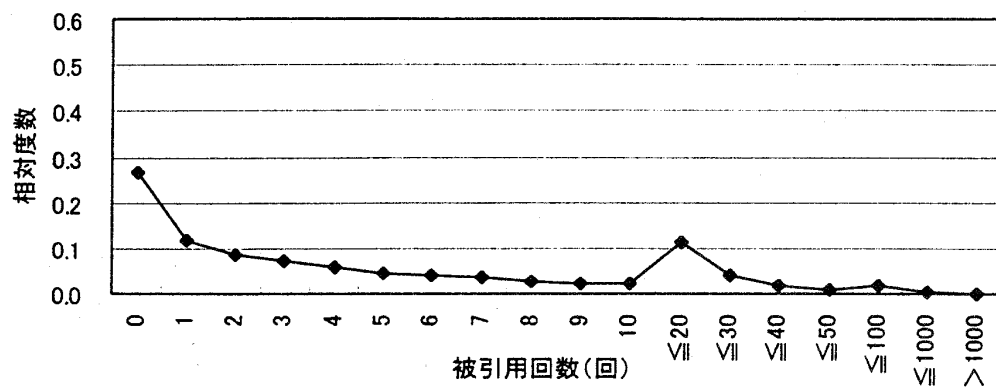


图1 相对被引用度数分布(全体)

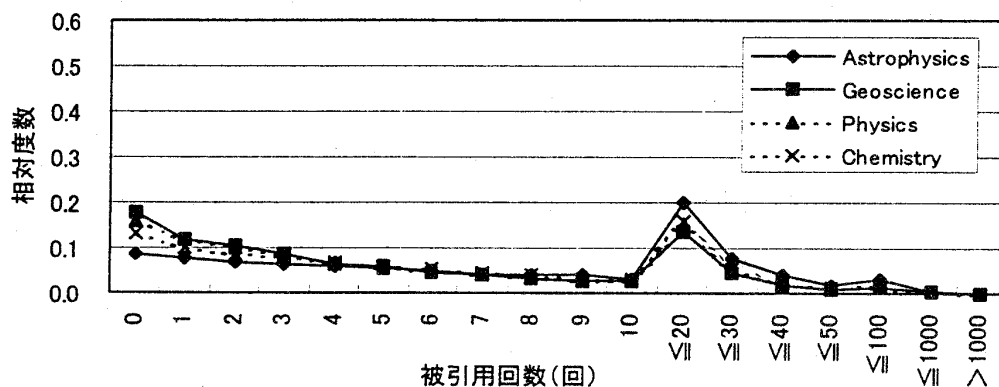


图2 相对引用度数分布(数学·物理系, 少死型)

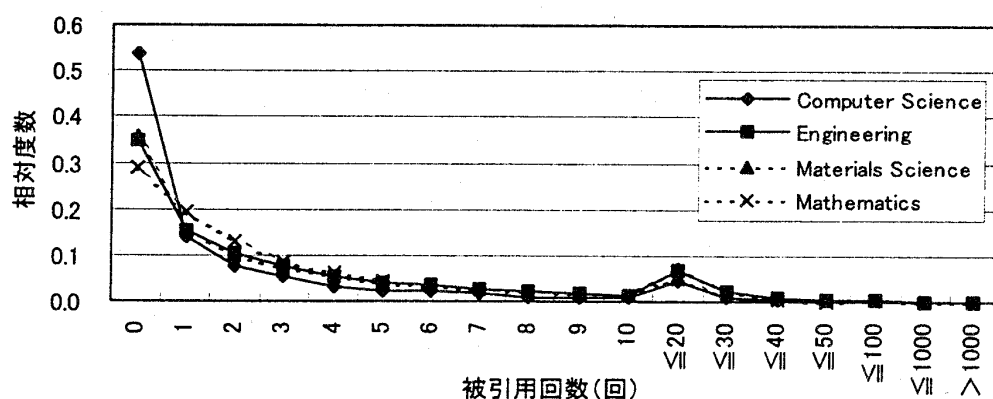


图3 相对被引用度数分布(数学·物理系, 多死型)

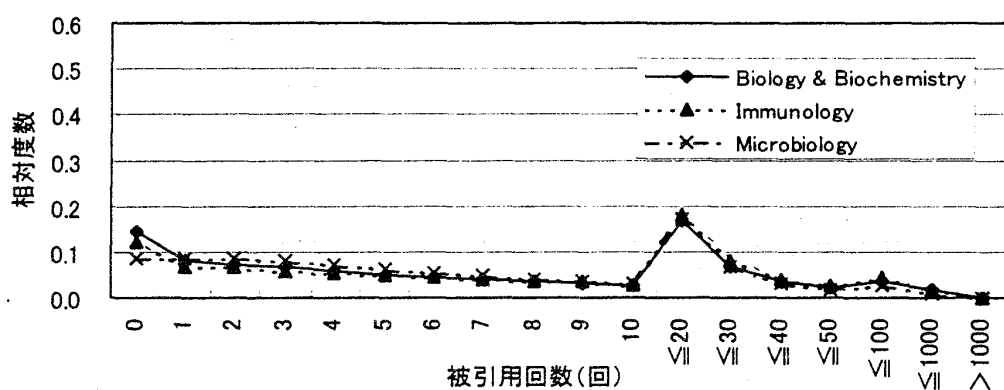


図4 相对被引用度数分布(生命科学系, 少死型)

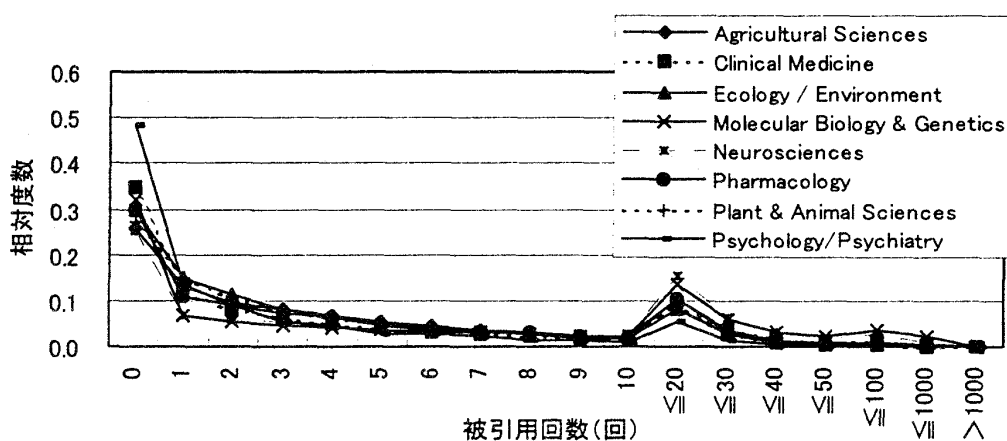


図5 相对被引用度数分布(生命科学系, 多死型)

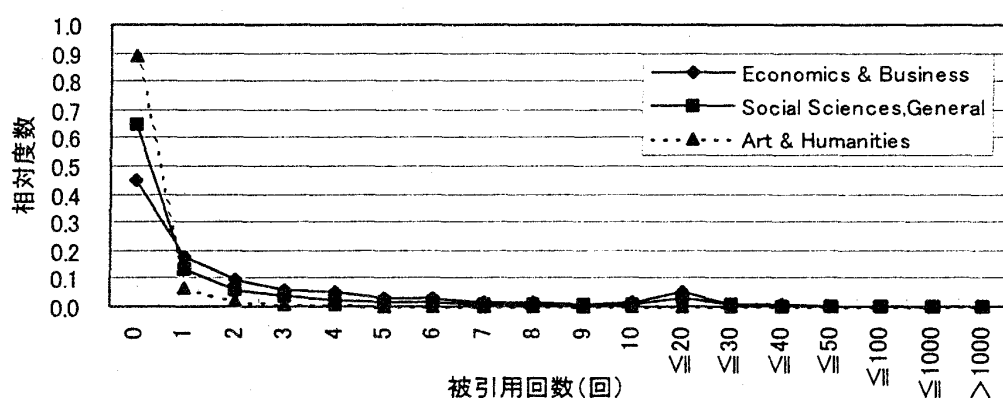


図6 相对被引用度数分布(人文・社会科学系, 多死型)

3. 3 被引用度の時間変化

81年から90年までの10年間の全体及び各分野の論文について被引用度の時間変化を図7～10に示す。全体においては、発行年には非引用度が0.21に過ぎないが、発行1年

後には 0.95、発行 2 年後には 1.27 に増加する。以下 1 年経過するごとに、1.20、1.10、0.99、0.90 と漸減する。

この傾向は個別分野においても同様であり、被引用度は数学・物理系と生命科学系は発行後 2 年ないし 3 年、人文・社会科学系では 4 年ないし 5 年で最大となり、その後漸減している。

数学・物理系では、Astrophysics、Physics、Chemistry の各分野は、被引用度のピークが顕著であり、逆に言えば陳腐化の速度も速い。一方、Geoscience、Computer Science、Mathematics の各分野は、被引用度のピークが顕著ではなく、古い論文が比較的引用されやすい分野と言えよう。特に Geoscience は、全論文の 80%以上が 1 回以上引用されている少死型の分野であり、被引用度 1 前後の水準を 5 年間維持していることは特筆に価する。

生命科学系では、Clinical Medicine における引用が顕著であり、発行 2 年後のピークにおいては、被引用度が 3 を越えている。

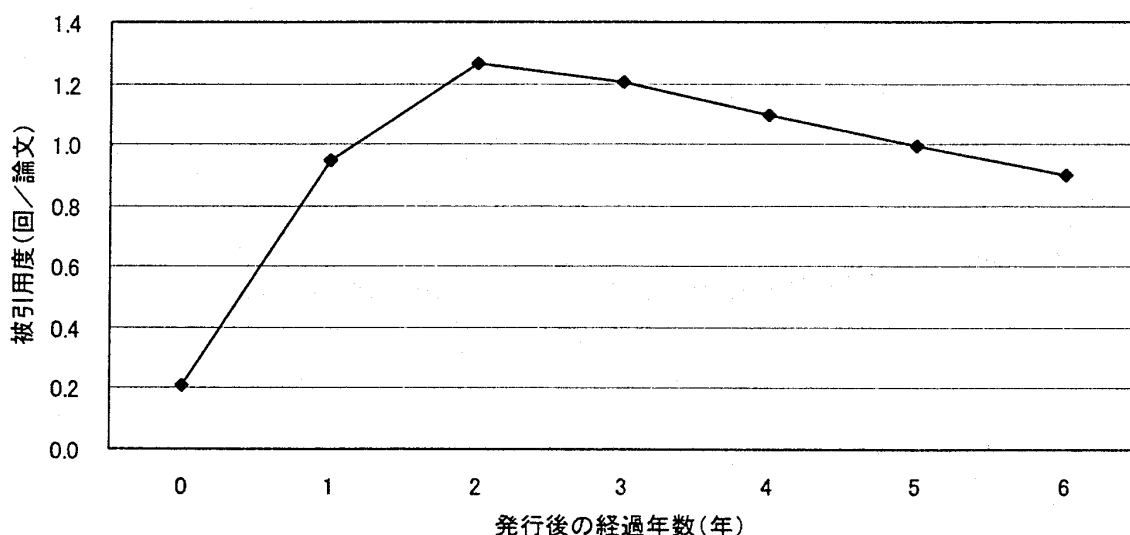


図7 発行後の被引用度変化(全体)

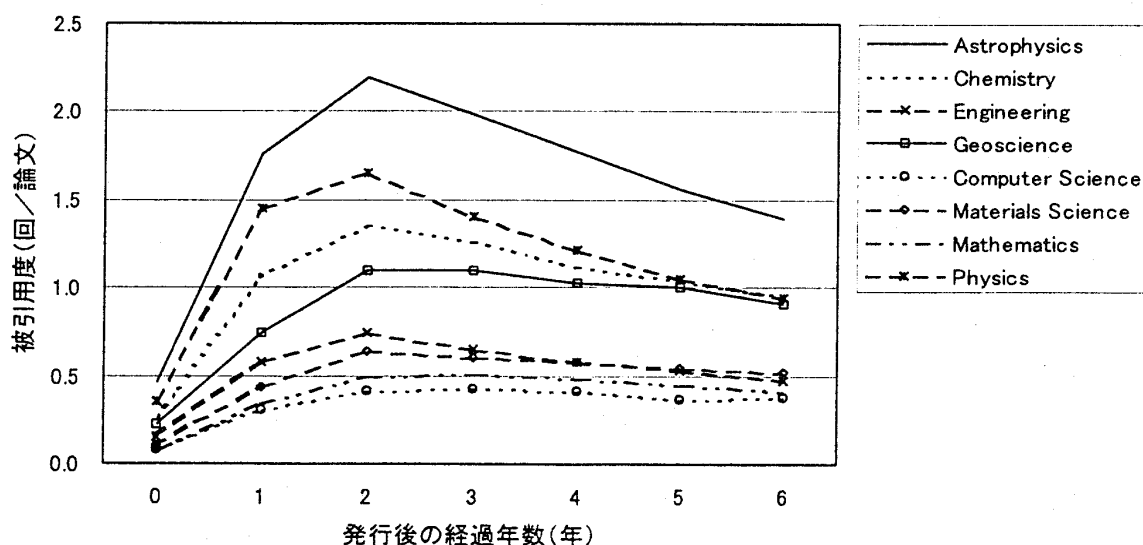


図8 発行後の引用度変化(数学・物理系)

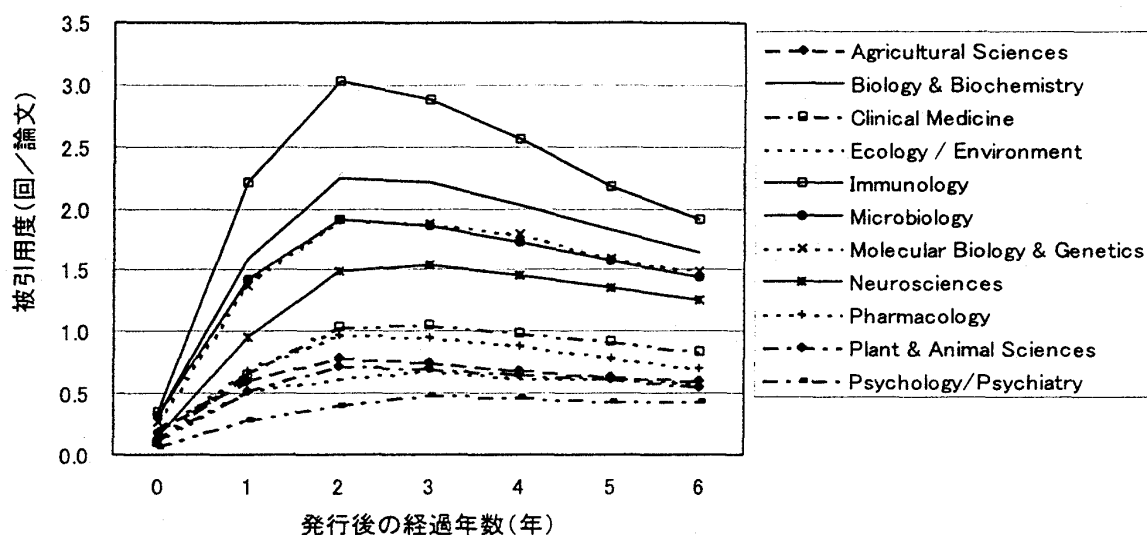


図9 発行後の引用度変化(生命科学系)

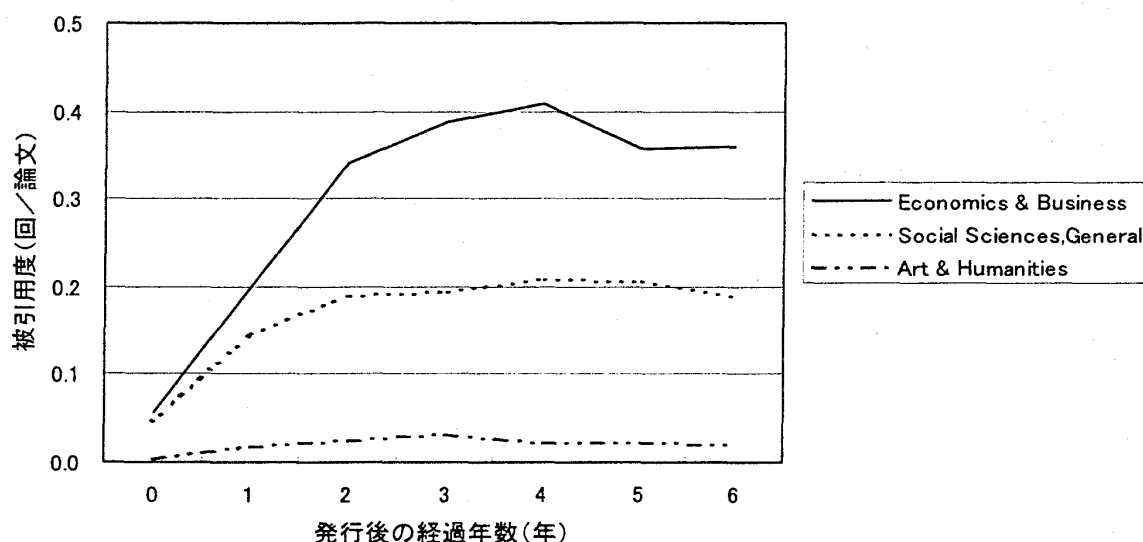


図10 発行後の引用度変化(人文・社会科学系)

3. 4 引用傾向による学問分野の類型化

学問分野のアクティビティを考える際に、その分野でどれだけの知識が産出されているか、さらに、産出された知識がどれだけ利用されているかが重要な指標となる。本節では、知識生産性の指標として相対論文数（全論文に占める当該学問分野の論文の比率）を、知識利用の指標として被引用度を取り、それらを組み合わせることにより、専門分野の特性を分析する（図 11）。ここで指標として用いる論文数は、あくまでも ISI の引用統計データベースに収録されている論文数であり、実際に流通している論文数とは異なる。しかしながら、ISI 社による収録対象としての評価が学界の需要をある程度の確に反映していると考えられるならば、「需要のある論文数」の指標として取り上げることができると考えられる。なお、相対論文数は、各分野で 50 倍ほどの開きが見られるため、分析の便からロジ

スティック変換した値を用いる。

全体的に Geoscience などの中庸な分野クラスタ（タイプ I）を中心として、左下に論文数・引用度ともに低い人文・社会科学系分野主体のクラスタ（タイプ II）が、上に被引用度が極めて大きい分野のクラスタ（タイプ III）が、右に被引用度は中程度で論文数の多い分野のクラスタ（タイプ IV）が位置している。タイプ II の分野は、いずれも多死型分野であり、ほとんどの論文は引用されないか、されても回数は多くない。タイプ III の分野は、引用が活発に行なわれているが、論文数はそれほど多くなく、もっとも多い Biology & Biochemistry でも 7 万件ほどである。このタイプの分野内においては、濃密な引用ネットワークが構成されているものと思われる。

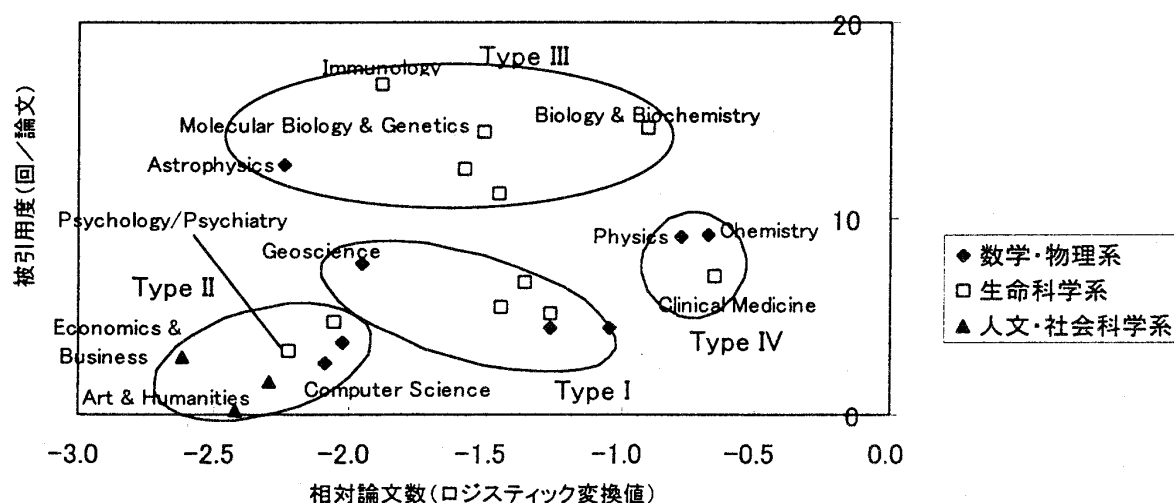


図11 各学問分野の論文数と引用度

4. まとめ

本研究では、我が国の学術論文における引用傾向を分析し、さらに論文数と引用度を元に学問分野の類型化を試みた。約 90 万件に及ぶ引用統計データから抽出することができる情報は膨大であり、本研究の分析から得た知見は各分野の引用傾向のごく一部に過ぎない。今後、同データを対象としたより緻密な分析手法を開発し、各分野における引用傾向の変化等について分析を進める必要がある。

参考文献

- 根岸 2000：根岸正光，孫媛，山下泰弘，西澤正己，柿沼澄男，「我が国の大学の論文数と引用数－ISI 引用統計データベースによる統計調査－」，学術月報，Vol.53，No.3，pp.64・80
- 孫 2000：孫媛，山下泰弘，西澤正己，柿沼澄男，根岸正光，「わが国各大学の論文数・引用数の状況(1981-1997)－ISI 引用索引データベースによる統計調査－」，学術情報センター

* 電気通信大学 University of Electro-Communications
+ 国立情報学研究所 National Institute of Informatics