

博士論文（要約）

論文の被引用数と引用持続性の間の  
相互関係、及びそれらと他の  
論文特性との関係に関する研究

小野寺 夏生

# 目 次

第1章 序論 .....	1
1.1 本研究の背景.....	1
1.1.1 研究評価への計量データの利用 .....	1
1.1.2 研究評価に計量書誌学データを用いることの意味と問題点.....	2
1.1.3 引用データの利用 .....	3
1.2 本研究の目的と意義.....	5
1.3 本論文の構成.....	7
1.3.1 第3章の概要：引用影響要因の可操作化と正確な測定の検討.....	7
1.3.2 第4章の概要：論文の被引用数に対する外在的要因の影響の検討.....	8
1.3.3 第5章の概要：論文の引用持続性を示す指標の性質の検討.....	9
1.4 本研究で用いたデータ .....	10
1.4.1 対象とする分野及びサンプル論文.....	10
1.4.2 被引用数データ及び外在的要因のデータ .....	13
第2章 先行研究のレビュー.....	15
2.1 論文の被引用数に及ぼす要因に関する研究.....	15
2.1.1 内在的要因による論文の被引用数への影響.....	15
2.1.2 外在的要因による論文の被引用数への影響.....	17
2.1.3 多くの潜在的影響要因を総合的に考慮した引用分析.....	26
2.1.4 引用影響要因とそれらを表す測度のまとめ .....	32
2.2 論文の引用履歴あるいは引用持続性に関する研究.....	33
2.2.1 論文の老化に関する数学モデルの開発 .....	33
2.2.2 高被引用論文(HCP)とそれ以外の論文の引用持続性比較.....	33
2.2.3 引用ウィンドウ期間による論文の引用ランクの変化.....	34

2.2.4 引用持続性による論文の分類及び引用持続性を測る指標 .....	35
2.2.5 引用持続性と論文の他の性質との関係 .....	36
2.3 論文の著者識別方法に関する研究 .....	37
第3章 引用影響要因の可操作化と正確な測定 .....	40
3.1 論文の被引用数に影響を与える要因とその可操作化 .....	40
3.1.1 内在的要因と外在的要因 .....	40
3.1.2 引用影響要因の可操作化 .....	42
3.2 本研究で用いる要因と測度の選択 .....	43
3.2.1 要因の選択 .....	43
3.2.2 各要因を表す測度の選択 .....	43
3.3 測度の正確な測定－著者の過去実績データの例 .....	46
3.3.1 著者論文数データ測定の概要 .....	47
3.3.2 著者識別対象の論文集合 .....	47
3.3.3 真論文と偽論文を判別する方法 .....	48
3.3.4 結果と考察 .....	50
3.3.5 まとめ .....	52
第4章 論文の被引用数に対する外在的要因の影響 .....	54
4.1 本章における研究目的 .....	54
4.2 データ源と方法 .....	56
4.2.1 対象とする分野及びサンプル論文 .....	56
4.2.2 被引用数データの取得 .....	56
4.2.3 被引用数に対する説明変数のデータの取得 .....	56
4.2.4 負の2項重回帰(NBMR)分析 .....	58
4.3 結果 .....	59
4.3.1 予備的な分析 .....	59

4.3.2 NBMR の結果.....	63
4.4 考察.....	71
4.4.1 被引用数に対する重要な影響要因.....	71
4.4.2 引用への著者のハロー効果はあるか?.....	72
4.4.3 説明変数間の相互作用.....	72
4.4.4 線形重回帰(LMR)分析と NBMR の適合度比較.....	72
4.4.5 サンプルングの問題.....	75
4.5 本章の結論.....	76
第5章 論文の引用持続性を示す指標の性質.....	77
5.1 本章における研究目的.....	77
5.2 データと方法.....	79
5.2.1 用いる持続性指標.....	79
5.2.2 用いるデータ.....	80
5.2.3 $D$ の算出.....	81
5.2.4 引用持続性と通算被引用数の間の関係.....	81
5.2.5 論文の諸特性から $D$ を説明する重回帰分析.....	82
5.3 結果.....	84
5.3.1 通算被引用数 $\alpha(T)$ の分布.....	84
5.3.2 $D$ の分布.....	87
5.3.3 $D$ と $\log[\alpha(T)+1]$ の間の関係.....	92
5.3.4 雑誌による引用持続性の違い.....	94
5.3.5 引用持続性と他の論文特性との関係に関する線形重回帰分析.....	95
5.4 考察.....	99
5.4.1 通算被引用数及びインパクトファクターと $D$ との関係.....	99
5.4.2 $D$ と他の論文特性の間の関係.....	101
5.4.3 将来の課題.....	103

5.5 本章の結論 .....	104
第6章 結論 .....	106
6.1 本研究で得られた主要な知見 .....	106
6.1.1 複数の分野に共通な傾向の確認 .....	106
6.1.2 $C$ と $D$ の関係 .....	107
6.1.3 $C$ 及び $D$ に影響を及ぼす特性の比較 .....	108
6.1.4 引用持続性指標の性質の分析 .....	108
6.2 本研究のオリジナリティ .....	109
6.2.1 研究全体にわたってのオリジナリティ .....	109
6.2.2 第3章の研究についてのオリジナリティ .....	110
6.2.3 第4章の研究についてのオリジナリティ .....	110
6.2.4 第5章の研究についてのオリジナリティ .....	110
6.3 本研究の限界 .....	111
6.3.1 内在的要因と外在的要因の交絡の可能性 .....	111
6.3.2 サンプル論文の発表年の限定について .....	111
6.3.3 著者実績測度に第一著者のデータのみを使用 .....	112
6.3.4 引用持続性の研究に15年間の引用ウィンドウのみを使用 .....	112
6.4 その他の考察 .....	112
6.4.1 ピアレビューによる論文評価との関係 .....	112
6.4.2 主題分野の設定について .....	113
参考文献 .....	115
謝辞 .....	125

## 目 次

図 4-1	各論文に対する $C_6$ の観測値( $C_{6i}$ )と NBMR による予測期待値( $\mu_i$ )の関係..68
図 4-2	各論文に対する $C_6$ の予測期待値( $\mu_i$ )と相対残差の関係.....69
図 4-3	NBMR の予測による被引用数確率分布と実際の分布との比較 .....70
図 4-4	LMR の予測による被引用数確率分布と実際の分布との比較.....74
図 5-1	Citation Delay ( $D$ )の意味を示す模式図.....80
図 5-2	$\log[\alpha T+1]$ の分布に対する Q-Q プロット ( $\alpha T \geq 1$ の論文を対象) .....86
図 5-3	$D$ の分布 ( $\alpha T \geq 5$ の論文を対象) .....87
図 5-4	$D$ の分布に対する Q-Q プロット ( $\alpha T \geq 5$ の論文を対象).....89
図 5-5	分野による分布範囲の違い (a) $D$ の分布 (b) $\log[\alpha T+1]$ の分布.....90
図 5-6	分野 A の[P25, P75]範囲に含まれる分野 B のパーセンタイル幅の分布.....91
図 5-7	$D$ の変化に対する $\log[\alpha T+1]$ の変化の非線形性.....93
図 5-8	$D$ と各説明変数間の Pearson 相関係数 $r$ .....96
図 5-9	消化器分野の 4 雑誌の間の $\alpha T$ - $D$ 分布の比較.....101
図 6-1	ピア評価のスコアと引用数スコアの関係を示す概念図.....113

## 表 目 次

表 1-1	選択した分野と雑誌.....	12
表 1-2	JIFによるクラス分けとサンプル誌の抽出.....	13
表 2-1	論文の質や内容に関する要因の引用への影響を扱った研究の概要.....	18
表 2-2	研究協力に関する要因の引用への影響を扱った研究の概要.....	20
表 2-3	著者に関する要因の引用への影響を扱った研究の概要.....	22
表 2-4	雑誌に関する要因の引用への影響を扱った研究の概要.....	23
表 2-5	参考文献に関する要因の引用への影響を扱った研究の概要.....	24
表 2-6	論文の可視性に関する要因の引用への影響を扱った研究の概要.....	25
表 2-7	種々の引用影響要因を総合的に考慮した研究の概要.....	31
表 2-8	引用影響要因とそれらを表す測度（先行研究のまとめ）.....	32
表 3-1	本研究で用いる引用影響要因の測度.....	46
表 4-1	目的変数( $C6$ , $C11$ )及び 14 の説明変数の平均値と標準偏差（括弧内）.....	60
表 4-2	$C6$ と各説明変数の間の Spearman 順位相関関数( $\rho$ ).....	61
表 4-3	説明変数の間の相関.....	62
表 4-4	4 分野以上で有意な相関を示す説明変数の組に対する Spearman 順位相関 係数.....	63
表 4-5	3 つの NMBR モデルの適合度比較.....	64
表 4-6	各説明変数の被引用数予測有意性.....	65
表 4-7	モデル C に対する NBMR の $x$ 標準化回帰係数 ( $s_j\beta_j$ ).....	66
表 4-8	論文被引用数予測力に関する種々の研究の比較(本章での研究を含む).....	67
表 4-9	平均平方相対残差( $MSRR$ )による NBMR と LMR の適合度比較.....	73
表 4-10	カイ二乗値( $\chi^2$ )とその有意確率( $p$ )による NBMR と LMR の適合度比較 ..	74
表 5-1	選定した主題分野、雑誌、及び抽出論文数.....	82
表 5-2	$D$ を予測する重回帰分析に用いる説明変数.....	83

表 5-3	重回帰分析のサンプルに用いた論文数.....	84
表 5-4	$\alpha(T)$ に関する統計.....	85
表 5-5	$\log[\alpha(T)+1]$ に関する統計.....	86
表 5-6	$D$ に対する統計量 ( $\alpha(T) \geq 1$ の論文を対象).....	88
表 5-7	$D$ に対する統計量 ( $\alpha(T) \geq 5$ の論文を対象).....	88
表 5-8	$D$ と $\log[\alpha(T)+1]$ の間の Pearson 相関係数 $r(D, O)$ .....	92
表 5-9	$D$ と $\log[\alpha(T)+1]$ の間の非線形な関係.....	93
表 5-10	各分野における雑誌の $D$ 平均値とインパクトファクター(JIF)の間の Pearson 相関係数( $r$ ).....	94
表 5-11	$D$ と各説明変数の間の Pearson 相関係数( $r$ ).....	95
表 5-12	説明変数の間の相関.....	97
表 5-13	重回帰分析の結果：決定係数、自由度調整済み決定係数、及び分散比.....	98
表 5-14	重回帰分析の結果：標準化偏回帰係数( $\beta$ ).....	99
表 5-15	引用持続性と種々の論文属性の間の相関に関する先行研究と本章の研究の 比較.....	102



# 第1章

## 序論

この章では、まず本研究の背景として、計量書誌学データ、特に論文の引用データを研究評価に利用することについての意味と問題点、及び利用に関して注意すべき事項について概観する。その注意事項のひとつに、論文の被引用数とその質や内容に直接関係のない要因(外在的要因)の影響を受けることがあるが、どのような外在的要因がどの程度論文の被引用数に影響するかについての知見を深めることが本研究の目的であることを述べる。その後、この研究目的に向けての3つの研究目標を設定し、それぞれの目標に対応した後段の各章の概要を記す。最後に本研究で共通に用いる論文データ、並びに被引用数及び外在的要因のデータについて説明する。

### 1.1 本研究の背景

#### 1.1.1 研究評価への計量データの利用

計量書誌学(bibliometrics)は、数学的・統計学的手法の適用による、情報流通過程における公表と伝達のパターンの研究である(Diodato, 1994)。種々の研究分野における研究動向の分析、研究コミュニティの構造・性質の解明等、研究を対象とした社会学的研究に、この手法が適用されてきた。その一環として、近年、論文、研究者、研究グループ、研究プロジェクトの評価への計量書誌学の利用が注目されている。

多くの研究評価で、論文の被引用数等の計量書誌学データが利用されている。Web of Science Core Collection (以下 Web of Science という)<sup>1</sup>や Scopus など引用索引を備えたデータベースの分析機能の向上、それらのデータに基づく InCites や SciVal などの情報分析ツール<sup>2</sup>の出現により、計量書誌学の専門家以外にとっても計量データの利用が容易になった。この利用が適切であれば、計量書誌学データは、専門家(ピア)による評価をより

---

<sup>1</sup> 従来の Web of Science データベースは、2015 年から Web of Science Core Collection と名を改め、Web of Science Core Collection その他の製品を提供する Thomson Reuters 社(現・Clarivate Analytics 社)のプラットフォームの名称が Web of Science となった。しかし、本研究の相当部分はこの改称以前に行われたため、本論文では、以前の名称である Web of Science またはその略称としての WoS を用いる。

<sup>2</sup> InCites は Clarivate Analytics 社、SciVal は Elsevier 社の提供する情報分析ツールで、それぞれ、Web of Science、Scopus のデータを情報源としている。

## 第1章 序論

妥当、公正にするための補完となり得る。しかしながら、データを補完材料として利用するのではなく、データに主導され、引きずられた評価が往々にして行われている。データや、それに基づいて計算される種々の計量的指標の意味や性質が十分理解されないまま、誤って利用される例もしばしばである。

計量書誌学データを研究評価に利用するには、その可能性と限界を理解して注意深く適用すること、そのために計量書誌学データ(特に引用データ)の性質をよく知ることが必要である。本研究の目的は、引用データの性質のある側面を解明することにあるが、この目的を述べる前に、研究評価に計量データを用いることの意味と問題点について若干の考察を行う。

### 1.1.2 研究評価に計量書誌学データを用いることの意味と問題点

#### 1.1.2.1 その意味

論文、研究者、研究組織に対する評価の視点は単一の軸によるものではなく多次的である。従ってその評価は、評価者の知識、経験、洞察、心証等から総合的になされるべきである。同分野の専門家(ピア)による評価はこの考えに立つものであり、今後とも研究評価の主たる方法であると考えられる。

このように、最終的評価は人間の主観によるものであるが、計量書誌学データはそれをより妥当、公正にするための補完材料として重要な意味を持ちうる。研究評価にこの種のデータを利用する意味は次のようにまとめられる。

- (1) 主観的判断に際して、一定の根拠あるいは裏付けを与えることができる。
- (2) ある種の評価視点(研究者の活動度、論文の学術的影響度など)に対しては適切な指標になり得る。
- (3) 偏った印象による評価や仲間うちの評価をチェックすることができる。

#### 1.1.2.2 その限界

まず、前述のように、計量書誌学データは人間による判断をより妥当、公正にするための補完材料であって、決して主観評価の代替になるものではない。この認識がないと、計量的指標の値を上げること自体が被評価者の目標になってしまう。

次に、これも上に述べたことであるが、評価の視点はいろいろあり、それぞれのデータはそれらのある視点を測るのみで、決して論文や研究者の質全般を測るものではない。たとえば研究者の発表論文数は主にその活動度に、論文の被引用数は主にその学術的影響度に関係する。つまり、一種類のデータですべてを決めることはできない。

もう一つの限界として、完全な計量データは一般に得られないということが挙げられる。個人や組織の論文発表のデータなら、完全な論文リストを入手することは可能だが、引用データは Web of Science や Scopus 等のデータベースに頼らざるを得ない。もちろんその

## 第1章 序論

カバレッジは完全ではなく、しかも、分野、発行国等による差が存在する。計量書誌データ、特に引用データの使用に際しては、常にこのことによるバイアスの可能性を考慮しなければならない。

計量書誌学データの適切な利用のあり方についてはこれまでしばしば論ぜられてきた(Leydesdorff, 2008; Glänzel and Moed, 2013)が、それらが結実した形で、最近、計量書誌学データの研究評価への利用に関するガイドラインと言うべき **Leiden Manifesto** が公表された(Hicks et al., 2015 ; 小野寺・伊神, 2016)。これは計量書誌学研究者によって書かれたものであるが、広く研究者、評価者、研究機関の管理者への普及を意図している。

### 1.1.2.3 データの誤った利用

計量書誌学データについての十分な理解がないまま、データや指標の値を過信することにより、データの誤った利用が生ずる。

典型的な例は、論文の被引用数を、その論文が発表された雑誌のインパクトファクター(JIF)と同一視する誤りである。JIFは、その雑誌の論文の平均被引用数であるから、当然個々の論文の被引用数とは異なる。しかも、論文の被引用数の分布は極めて歪度が大きい(skew)ので、平均値であるJIFは、その雑誌の被引用数の代表値としては適切と言えない(Seglen, 1994)。JIFの誤用に対し、たとえば、2013年5月に出された **San Francisco Declaration on Research Assessment (DORA)**では、JIFを論文、研究者個人、研究組織等の評価に用いることを厳しく戒めている(The American Society for Cell Biology, 2013)。この宣言には11,671名の個人(学会誌編集者等)と447の団体(学会等)が署名している(2018年2月12日現在)ので、このような考えは、学界にもある程度浸透してきたと見られる。

別のデータ誤用として、評価指標の計測値による研究組織や研究機関のランク付けがある。計量書誌学指標の値はいろいろな要素に影響され、不確定性(偶然的及び系統的な誤差)を持つので、その僅かな差には意味がない。しかし、大学ランキング等では、意味のない僅かな評価値の差により、事実上有意差のない大学の間にも順位差が付けられている(Bornmann et al., 2013)。

### 1.1.3 引用データの利用

研究評価に利用される計量書誌データのうち特に重要なのは論文の引用データである。被引用数は、論文の学術的影響度(インパクト)の定量的尺度として使われるだけでなく、それに基づいて、研究者(論文の著者)、研究グループ、研究機関等の研究インパクトを測るための種々の指標が提案されている(Waltman, 2016)。引用データの研究評価への利用に関する枠組み、方法、得られる情報と問題点、事例等については、Moedの著書に詳しく論ぜられている(Moed, 2005)。

## 第1章 序論

1.1.2 で述べた計量書誌学データ利用の意味と問題点は、引用データにもほぼ当てはまるが、ここでは、主に引用データに特有のことからについて考察する。

### 1.1.3.1 引用データを研究評価に用いることに対する批判

これについては、昔から次のような批判がある(Lindsey, 1989; MacRoberts and MacRoberts, 1987, 1989, 1996, 2010)。

- (1) 引用をする理由、動機は、引用者により、また個々の引用によりさまざまなので、それを単純に集計することによりのような意味があるか疑問である。
- (2) 本質的・重要な引用もあればおざりな引用もあり、また、被引用論文の記述を否定したり批判したりする引用もある。これらを一律に数えることはおかしい。
- (3) 論文が主要な研究成果発表媒体でない分野が存在する。
- (4) 質の高い研究が多く引用を得るとは限らない。
- (5) 技術的論文が設計や製造に利用される等、引用以外での利用が多い論文がある。

### 1.1.3.2 引用データを研究評価に用いることの妥当性

上記の批判にはいずれも相当の事実が含まれる。しかし、注意して使用すれば、引用データから有用な情報が得られるのも事実である。研究評価への利用の観点から見ての引用データの有用性は、次の2点によると考えられる。

- (a) 論文の学術的影響度を測るには、引用データには一定の妥当性があり、現時点で考える限り最も適切な計量データである。
- (b) かなり網羅的で信頼性の高いデータ源となるデータベース(Web of Science, Scopus)が存在する。

論文の影響度(インパクト)を測る計量データとして、最近、オンライン閲読データ、オルトメトリクスデータも注目されている。前者は、電子ジャーナルサイトや文献管理サイト(Mendeley など)におけるダウンロード数から得られ、後者は種々のソーシャルネットワークサービス(SNS)における論文への言及数から得られる。いずれも計量書誌学データではないが、引用とは別の意味での影響度(学術的より社会的)を示すデータとして発展の可能性があるかもしれない。しかし、これらのデータには次のような問題があり、引用データに比べて信頼性が低いと考えられる(Gumpenberger et al., 2016)。

- (a) 論文のダウンロードや SNS 言及は、その利用とは必ずしも結びつかない。引用においても、中身を読まずに書誌リストからコピーすることもあるが、それでも何らかの関係を認めてのことであろう。
- (b) ダウンロードや SNS 言及のほとんどは、論文の公開直後になされることが知られており、長期的な傾向を知ることはできない。
- (c) 特に SNS 言及の場合、その計量が何を測定しているのか曖昧である。
- (d) 自分の論文を何回もダウンロード・言及するなど、アンフェアな操作が容易である。

### 1.1.3.3 引用データ利用で注意すべきことがら

引用データを学術情報評価あるいは研究評価に利用するに当たっては、次のことを十分認識して注意深く取り扱うことが必要である。

- (1) 引用指標自体が評価対象ではなく、ピアによる主観的評価を補完・強化する材料である。
- (2) 引用データは論文の質のある側面(学術的影響度・注目度)にのみ関係する。従って、他のデータ・情報と併用する必要がある。
- (3) 1.1.3.1 に示された批判には、相当の事実が含まれることを考慮すべきである。
- (4) 論文の被引用数や、それから求められる種々のインパクト指標は、極めて歪度の大きい(skew)分布を示す(Seglen, 1992)。つまり、一部の論文や非評価者が高い値を持ち、大多数は低い値に集中しているため、それに基づいて細かいランク付けをすることは不適當である。
- (5) 論文の被引用数は、その質や内容に直接関係のない要因(外在的要因)の影響を受ける<sup>3</sup>。

論文の被引用数が、その重要性の尺度となることは統計学的には受け入れられるとしても、被引用数(あるいはそれに基づく指標)とピア評価の結果は必ずしも一致しない。多くの研究において、この間の相関は「ほどほど」(相関係数 0.4~0.6 程度)であることが示されている(Oppenheim, 1997; Rinia et al., 1998; 林, 2003; Abramo et al, 2011b, 2013; Mryglod et al, 2013a, 2013b)。このことは、上記の(2)~(5)を考えれば当然予想されることである。

## 1.2 本研究の目的と意義

本研究の目的は、1.1.3.3 で述べた(5)に関係する。すなわち、どのような「外在的」要因がどの程度論文の被引用数に影響するかについての知見を深めることにより、引用データの利用に資する情報を与えることを意図する。

この目的に向けて、本研究では、以下の3つの研究目標を設定し、この順序で考究を進める。

研究目標 1：引用影響要因の可操作化と正確な測定の検討

---

<sup>3</sup> 論文の被引用数に影響を与える可能性のある要因のうち、その質や内容に関する要因を「内在的」要因(“intrinsic” factors)、質や内容には直接関係のない要因(たとえば著者の数や名声、参考文献数など)を「外在的」要因(“extrinsic” factors)という(Chen, 2012; Peng and Zhu, 2012)。

## 第1章 序論

論文の被引用数に影響を与える可能性のある外在的要因には様々なものがあり、かつ、それらの要因をどのような測度を用いて測定するか（すなわち、要因という概念の可操作化）にも選択の幅がある。ここではまず、本研究の目的に照らして適切な外在的要因とその測度の選択を行う。更に、これらの中で最も測定が難しい論文著者の過去実績に関する測度について、比較的簡便に正確な測度データを得る方法を開発する。

### 研究目標 2：論文の被引用数に対する外在的要因の影響の検討

論文の被引用数に影響を及ぼす種々の外在的要因の影響を体系的に検討し、どの要因がどの程度の影響力を持つか分析する。複数の分野についてこれを行うことにより、影響を与える主要因について分野共通の傾向があるかどうかを明らかにする。

### 研究目標 3：論文の引用持続性を示す指標の性質の検討

論文の被引用数を観測する期間（引用計測期間）は、被引用数に影響を及ぼす外在的要因の一つであるが、この影響の検討は、研究目標 2 とは別途の方法によることが適当である。このため、論文の引用持続性を示す指標の分布の特徴、長期的被引用数との関係、外在的要因への依存性を、複数の分野において検討する。

研究目標 1、研究目標 2、研究目標 3 に対応した研究は、それぞれ第 3 章、第 4 章、第 5 章で述べる。

本研究のオリジナルと考えられる点を以下に挙げる。

#### (1) 研究全体に関して

- ・研究目標 2 と研究目標 3 で述べたことを、複数の分野において実施することにより、論文の被引用数に影響を及ぼす外在的要因、論文の引用持続性を示す指標の性質について、分野を超えた一般的傾向を見出すことができるか検証する。
- ・研究目標 1 において、論文の被引用数に影響を与える外在的要因を体系的に整理する。その結果を研究目標 2 及び 3 の研究に適用して、各要因の影響を体系的に分析する。

#### (2) 研究目標 1 に関して

- ・論文著者の過去実績に関する測度である著者の過去論文数を正確に得るために、同名異人著者の論文を識別する比較的簡便な方法を開発する。

#### (3) 研究目標 2 に関して

- ・外在的要因の被引用数への影響の分析において、観測期間の異なる 2 通りの被引用数を目的変数とすることにより、安定した結果が得られるかどうか検証する。
- ・被引用数に影響を及ぼす外在的要因について、異なる分野に共通の傾向が見られるかどうか検討する。

#### (4) 研究目標 3 に関して

- ・論文の被引用数を示す指標の性質はほとんど明らかにされていないので、最も適切と考えられる指標を選んで、その基本的性質（分布の形状、長期被引用数との関係など）を解明する。

## 第1章 序論

- ・引用持続性に影響を与える外在的要因を見出し、研究目標 2 で見出した被引用数に影響を与える要因と比較する。

これらのオリジナルな点が実際に達成されたかどうかについては、第 6 章で述べる。

### 1.3 本論文の構成

この第 1 章は、研究の背景と目的、及び本論文の構成について述べるとともに、第 3 章～第 5 章で行うデータ分析に共通に用いるデータについて説明する。第 2 章では、第 3 章～第 5 章の研究に関する先行研究のレビューを行う。

第 3 章、第 4 章、第 5 章はそれぞれ、1.2 に示した研究目標 1、研究目標 2、研究目標 3 に対応した研究について述べる。そして第 6 章で、全体的考察を行い、結論を述べる。

第 3 章第 3 節、第 4 章、第 5 章の内容は、学術雑誌に発表された以下の 3 つの論文が核となっている。

- ・第 3 章第 3 節の核となっている論文：

Natsuo Onodera, Mariko Iwasawa, Nobuyuki Midorikawa, Fuyuki Yoshikane, Kou Amano, Yutaka Ootani, Tadashi Kodama, Yasuhiko Kiyama, Hiroyuki Tsunoda and Shizuka Yamazaki. A method for eliminating articles by homonymous authors from the large number of articles retrieved by author search. *Journal of the American Society for Information Science and Technology*, 2011, vol. 62, no. 4, p. 677–690. DOI: <http://doi.org/10.1002/asi.21491>

- ・第 4 章の核となっている論文：

Natsuo Onodera and Fuyuki Yoshikane. Factors Affecting Citation Rates of Research Articles. *Journal of the Association for Information Science and Technology*, 2015, vol. 66, no. 4, p. 739–764. DOI: <http://doi.org/10.1002/asi.23209>

- ・第 5 章の核となっている論文：

Natsuo Onodera. Properties of an index of citation durability of an article. *Journal of Informetrics*, 2016, vol. 10, no. 4, p. 981–1004. DOI: <http://doi.org/10.1016/j.joi.2016.07.001>

以下に、第 3 章、第 4 章、第 5 章の概要を記す。

#### 1.3.1 第 3 章の概要：引用影響要因の可操作化と正確な測定の検討

論文の被引用数とその学術的影響度の尺度であるとすれば、それに基本的に影響するのは論文の質や内容に関わる「内在的」要因であろう。しかし、これに対して、質や内容に

## 第1章 序論

直接関係はないが、被引用数に影響を与える可能性のある種々の「外在的」要因も存在する。

被引用数に対する各要因の影響を定量的に検討するには、これらの要因を測定可能な測度(変数)で表現する(可操作化)必要があるが、一つの要因に対して何を測度にするかについて、いろいろな選択肢があり得る。論文の被引用数に影響を及ぼす要因、及びそれらの要因を表す測度については多くの研究<sup>4</sup>があるが、それらの結果を踏まえ、本研究の目的に照らして、被引用数への影響を検討するための外在的要因とそれらを表す測度を選択する。

次に、選択した測度の中で最も測定が困難(要する労力、エラーの生じ易さの両面で)である論文著者が過去に発表した論文数について、実際のデータを用いて、比較的簡便に正確なデータを得る方法を提案し、得られた結果を検証する。著者の過去論文数を調べる一般的な方法は、データベースを用いて著者名による文献検索を行うことであるが、その際、同姓同名の異なる著者の存在が大きな問題になる(Moed, 2005, Chapter 14)。本研究で用いた論文データにおいては、約2,500人の著者に対する著者名サーチで検索された論文は60万件以上にのぼった。これらを真の著者による論文(真論文)と同名異人著者による論文(偽論文)に手作業で分けることは不可能であるため、半自動的に判別する方法を検討した。この種の著者識別方法は種々研究・開発されている<sup>5</sup>が、ここでは、同定洩れ(真論文を偽論文と判別)を多少犠牲にしてもノイズ(偽論文を真論文と判別)を最小限に抑えることを目標に、より簡便な方法を考案した。

### 1.3.2 第4章の概要：論文の被引用数に対する外在的要因の影響の検討

論文の被引用数に影響を与えるいろいろな外在的要因について、1.3.1に述べたように多くの研究がなされている。しかしながら、どの要因が被引用数に有意な影響を及ぼすか、その影響の強さはどの程度か、について明確な合意は得られていない。その理由の一つは、これまでの研究の多くが、ある単一の要因に着目している(あるいは複数の要因をそれぞれ独立に見ている)ために、要因間の相互作用が明らかでないことにある。もう一つの理由は、様々な潜在要因を総合的に考慮したいいくつかの研究(多くは重回帰モデルが使われる)は、対象の論文集合が特定の分野や雑誌に限定されているため、結論の一般性に限界があることである。

第4章では、第3章で選択した要因を可操作化した測度のデータを用いて、それぞれの外在的要因が論文の被引用数に及ぼす影響を研究する。第3章で述べるように、論文の被引用数が属する主題分野、論文のタイプ、言語に大きく依存することは明らかになっているので、いくつかの異なる分野について、同じ年にいくつかの雑誌に発表された同タイプ・同言語の論文(英語の原著論文)の間の被引用数の違いと、種々の要因へのその依存性を

<sup>4</sup> これらの研究については2.1において詳述する。

<sup>5</sup> これらの研究については2.3において詳述する。



## 第1章 序論

体系的に検討する。同じ年に発表された同じ分野、同じ論文タイプ、同じ言語の論文の中で被引用数を比較することにより、これらの影響は除去され、それ以外の外的要因の影響を分析することができる。また、複数の分野についてこの分析を行うことにより、影響を与える主要因について分野共通の傾向があるかどうかを明らかにすることができる。

被引用数への種々の要因の正味の寄与（他の要因の影響を除去した寄与）を評価するために負の2項重回帰分析を用いる。これらの要因には論文の質や内容に関する特徴は含まないので、得られるモデルは被引用数をある程度予測できたとしても、高精度の予測は期待できない。この研究の目標は説明力の高いモデルを得ることではなく、外的な要因から期待される被引用数のベースラインを得ることにある。分野を越えてある程度共通性の高いベースラインを得ることができれば、実際の論文のこのベースラインからの乖離は、その論文のインパクトをこれまでより適切に測る指標となり、研究評価に引用データがより正確に応用できるかもしれない。

### 1.3.3 第5章の概要：論文の引用持続性を示す指標の性質の検討

被引用数が同じ論文でも、その経時分布は一樣ではない。多くの論文は、発表後しばらくはあまり引用されず、次第に注目されて発表後2～6年後に引用のピークに達し、その後減衰していくというパターンをとる。しかし、発表後1～2年間にその大部分が引用されてしまう論文もあるし、長期間コンスタントに引用される論文もある。あるいは、数年間は全く引用されず、10年以上経って突然引用されるようになるものもある。

論文の被引用数が時間経過とともにどのように変化するか、つまり論文の引用の老化(aging または obsolescence)あるいは持続性(durability)については、いろいろな観点から研究がなされている<sup>6</sup>。しかし、論文の被引用数自体についての研究に比べると、引用持続性については系統的な知見の蓄積が不十分である。その本質的理由の一つは、論文の引用持続性を測る定量的な指標が確立されていないことにあると考えられる。

従来からよく使われている遅延型(delayed または sleeping beauty)、早期型(flash in the pan)、正規型(normal)の分類法(Garfield, 1980; Glänzel et al., 2003; van Raan, 2004; van Dalen and Henkens, 2005; Costas et al., 2010; Li and Shi, 2016)は、その定義や基準が人により異なる。これに対して Wang et al. (2015)により導入された Citation Delay は、被引用年齢分布の全域の情報をういた測度であり、引用持続性を示す測度としてこれまで提案されたものの中で最適と考えられる。しかし、その性質についてはまだほとんど解明されていない。また、論文の引用持続性と論文が持つその他の特性の関係を系統的に比較した研究も多くない。

第5章では、この Citation Delay を論文の引用持続性指標として使い、以下の3

---

<sup>6</sup> これらの研究については2.2において詳述する。

## 第1章 序論

項を研究目的とする。

- (1) Citation Delay の分布の特徴を検討する。
- (2) 異なる分野において、この持続性指標と論文の被引用数の関係を調査し、早期に引用される論文と遅延的に引用される論文の間で、長期的被引用数に傾向の違いがあるか否かを検討する。
- (3) 異なる分野において、論文の諸特性（外在的要因）と引用持続性の関係を調べ、分野を超えた共通の傾向が見られるか否かを検討する。

### 1.4 本研究で用いたデータ

第3章、第4章、第5章で行うデータ分析では、共通の論文をサンプルとして用いている。また、それらの論文に対して取得した被引用数データ及び外在的要因のデータもかなり共通している。それらについてここで説明する。

#### 1.4.1 対象とする分野及びサンプル論文

1.2 で述べた研究目的に基づいて、本研究では、いくつかの異なる分野について、同じ年に発表された同じタイプの論文（英語の原著論文）の間で、被引用数及び引用持続性の違い、それらに及ぼす種々の外在的要因の影響を分析する。異なる分野、論文タイプ、言語の論文の間では、引用頻度その他の論文の性質に傾向の違いがあることは既に明らかになっているので、サンプルをこのように限定することにより、これらの影響を除去し、それ以外の外在的要因の影響を分析することを意図した。また、複数の分野についてこの分析を行うことにより、影響を与える主要因について分野共通の傾向があるかどうかを明らかにすることができる。

この考えの下に、対象とする分野、雑誌、論文を以下のように選定した<sup>7</sup>。

##### 1.4.1.1 分野の選定

Web of Science のすべての主題カテゴリー (Subject Categories) を 11 の広域分野に分け、そのうち雑誌数の多い 6 広域分野（物理学、化学、工学、生物学、基礎医学、臨床医学）を同定した。そして、これらの各広域分野から、(a) その広域分野を代表し得る、(b) 年間総論文数が 5,000 以上、(c) 論文数で重み付けしたカテゴリー内 JIF 平均値が広域分野

---

<sup>7</sup> この細節の以下に示す Web of Science の主題カテゴリーとそれらに属する雑誌、それらの雑誌の JIF 等のデータは、Journal Citation Reports Science Edition 2004 に基づく。本研究では 2000 年に発表された論文がその後 6～14 年の期間に得た引用のデータを扱うため、中間的時期である 2004 年を選んだ。

## 第1章 序論

の全平均 JIF に近い、ことを基準として、主題カテゴリーを一つずつ選んで調査対象とすることとした。選択した6つの主題カテゴリーは次の通りである。

- ・物理学：物性物理学(condensed-matter physics)
- ・化学：無機・核化学(inorganic and nuclear chemistry)
- ・工学：電気・電子工学(electric and electronic engineering)
- ・生物学：生化学・分子生物学(biochemistry and molecular biology)
- ・基礎医学：生理学(physiology)
- ・臨床医学：消化器病学(gastroenterology)

以下では、これらの主題カテゴリーを「分野」と呼び、分野名の下線を引いた部分を略称として用いる。

### 1.4.1.2 各分野における雑誌の選定

6つの分野のそれぞれから4誌ずつ、計24誌を選んだ。これらを表1-1に示す。この選定には以下の点を考慮した。

- (a) その主題カテゴリーのみに分類されている雑誌から選ぶ（Web of Scienceでは、収録される雑誌の50%近くが2つ以上の主題カテゴリーに分類されている）。
- (b) 英語のみを用いている雑誌とする。
- (c) 1分野の中に複数の発行国の雑誌を含む。
- (d) JIFが高い雑誌とあまり高くない雑誌を含む。

雑誌を(a)のように限定したのは、その分野のできるだけ純粋な傾向を把握するためである。同じ主題カテゴリーの中でも、サブ領域により引用頻度等の論文の性質には傾向の違いがあることが知られている(Moed, 1989; Ruiz-Castillo and Waltman, 2015)。他の主題カテゴリーにも分類されている雑誌は、当該の分野とはやや異質のサブ領域を形成していると考えられる。

(d)のようにしたのは、その分野の全論文を代表するサンプルとするためである。これについては、次の1.4.1.3で更に詳しく説明する。

表 1-1 選択した分野と雑誌

分野	略称	雑誌名 <sup>a</sup>	出版国	論文数 <sup>b</sup>	サンプル論文数
物性物理学	物性	European Physical Journal B	DEU	538	60
		Journal of Physics – Condensed Matter	GBR	813	56
		Physica B	NLD	148	59
		Physical Review B	USA	4738	55
無機・核化学	無機	Inorganic Chemistry	USA	931	53
		Inorganica Chimica Acta	CHE	546	60
		Journal of the Chemical Society – Dalton Transactions	GBR	682	54
		Transition Metal Chemistry	NLD	139	60
電気・電子工学	電気	IEE Proceedings – Circuits, Devices and Systems	GBR	52	51
		IEEE Transactions on Circuits and Systems I – Fundamental Theories and Applications	USA	218	60
		IEEE Transactions on Microwave Theory and Techniques	USA	295	59
		Signal Processing	NLD	178	59
生化学・分子生物学	生化	European Journal of Biochemistry	GBR	788	60
		Journal of Biochemistry (Tokyo)	JPN	275	60
		Journal of Biological Chemistry	USA	5504	60
		Journal of Molecular Biology	USA	875	60
生理学	生理	Journal of General Physiology	USA	110	60
		Japanese Journal of Physiology	JPN	72	60
		Journal of Physiology – London	GBR	472	58
		Pflugers Archive European Journal of Physiology	DEU	238	58
消化器病学	消化器	American Journal of Gastroenterology	USA	430	58
		Gastroenterology	USA	259	59
		Gut	GBR	277	56
		Journal of Gastroenterology	JPN	124	60

a) 2000年当時の誌名（その後変更された名称あり）

b) 2000年に発表された“article”の数

#### 1.4.1.3 分析対象論文の選択

こうして選んだ24誌から、2000年に発表された正規論文（Web of Scienceの記事タイプが“article”であるもの）を分析の対象とし、Web of Scienceを用いてそれらの論文を特定した。但し、2ページ以下の論文、会議プロシーディングス論文（“article”と同時に“proceedings paper”にも分類されているもの）、及び著者または所属アドレスのデータが欠落している論文を除外した。

第4章、第5章で主要な分析方法とした重回帰分析においては、これらの各雑誌からそれぞれ50-60件を無作為抽出したものをサンプルとした<sup>8</sup>。表1-1には、各雑誌の2000年発表正規論文数、それから抽出したサンプル論文数も示す。

<sup>8</sup> 第5章の重回帰分析以外の分析（引用持続性指標の分布、同指標と長期被引用数の関係等）では、各雑誌の全正規論文を用いた。

## 第1章 序論

表 1-1 に見るように、論文数は雑誌により大きな差があるにも拘わらず、重回帰分析のためのサンプル論文の数はほぼ同数である。このように、規模が異なる雑誌からほぼ同数の論文を抽出した理由は次の通りである。

サンプルはそれぞれの分野を代表しなければならないので、その分野のインパクトの高い雑誌と比較的低い雑誌からサンプリングを行う必要がある。そこで、各分野の雑誌（当該の主題カテゴリーのみに分類された雑誌）を、JIF に従って論文数がほぼ均等になるように4つのクラスに分け、各クラスから、1.4.1.2 で述べた点を考慮して1誌ずつを選んだ。たとえば、無機分野では表 1-2 のようになる。

従って、各雑誌はほぼ論文数が等しい4つのJIFクラスのそれぞれを代表するので、各誌からほぼ同数の論文を抽出することが妥当である。

表 1-2 JIF によるクラス分けとサンプル誌の抽出（無機分野の場合）

JIF の範囲	雑誌数	論文数の比率	サンプルとした雑誌	
			誌名	JIF
> 3	2	23.1%	Inorg Chem	3.45
2-3	2	22.0%	Dalton Trans	2.93
1-2	3	25.0%	Inorg Chim Acta	1.55
< 1	7	29.9%	Trans Met Chem	0.86

しかし、物性分野と生化学分野では、それぞれ Physical Review B と Journal of Biological Chemistry がその分野の論文数のかなりのシェアを占めており（それぞれ 45%、44%）、上記の配分がうまくできなかった。つまり、この2つの分野では、論文数が多くインパクトも高いこの2つの雑誌の寄与が過小評価され、バイアスが生じている可能性がある。重回帰分析では、雑誌をダミー変数とする、あるいは雑誌の JIF を変数とすることで、この問題の緩和を図った。

### 1.4.2 被引用数データ及び外在的要因のデータ

1.4.1 で述べた考えに従って選択した雑誌の論文に対して、分析に用いるデータを次のように取得した。

#### 1.4.2.1 被引用数データの取得

第4章で述べる研究（被引用数に及ぼす諸要因の影響）に用いる被引用数データは、2006年10月と2011年12月に Web of Science を用いて取得した。論文の発表年はすべて2000

## 第1章 序論

年なので、引用ウィンドウの長さはそれぞれ6-7年、11-12年になる。

第5章で述べる研究（引用持続性指標の性質）に用いる被引用数データは、やはり Web of Science から 2015 年 3 月 31 日に取得し、2000 年から 2014 年までの各年の被引用数を記録した。

これらの被引用数には自己引用（引用論文と被引用論文の間に共通の著者が存在する引用）を含むが、これについて触れておきたい。被引用数の多い論文には自己引用がかなり寄与しているのではないか、あるいは、多数の共著者による論文がよく引用されるのはそれぞれの著者がその論文を自己引用する機会が多くなるからではないか等の疑問が生ずるかもしれない。これらの疑問が是認されるならば、論文の引用インパクトに影響する要因を分析する際には自己引用を除く必要がある。しかし、Aksnes (2003a)や Glänzel et al. (2004)による研究では、長期にわたる大量のデータの分析から、論文の被引用数が増えるに連れて他者被引用に対する自己被引用の率は減る傾向にあることが示された。また、Glänzel and Thijs (2004)及び Aksnes (2003a)は、著者数の増とともに自己引用率は減少することを示した。これらの結果から、少なくとも、本研究のような巨視的な視点からの統計解析では、自己引用を除く必要はないと考えられる。発表後短期間の引用では自己引用のバイアスが大きいことが示されている(Aksnes, 2003a)が、本研究で用いた引用ウィンドウは比較的長期（発表後少なくとも6年）なので、その影響はほとんどないと考えられる<sup>9</sup>。

### 1.4.2.2 被引用数に影響の可能性を持つ外在的要因のデータの取得

第4章と第5章では、それぞれ論文の被引用数、引用持続性に対する種々の外在的要因の影響を検討するため重回帰分析を用いる。そこで用いられる説明変数(影響要因を表す測定)の選択については第3章で説明し、データ取得と数値化の詳細については第4章、第5章で述べる。

---

<sup>9</sup> 自己引用を除いたデータによる分析も行ったが、結果に影響を及ぼす違いは見出せられなかった。

## 第2章

### 先行研究のレビュー

本章では、本研究に関する先行研究のレビューを行う。2.1、2.2、2.3は、それぞれ、第4章（論文の被引用数に対する外在的要因の影響）、第5章（論文の引用持続性を示す指標の性質）、第3章第3節（著者の過去論文数データの正確な測定）に関するものである。第3章第3節に関するレビューを最後に置いたのは、この関係の研究は本研究の主要目的（論文の被引用数及び引用持続性と外在的要因の関係）からやや離れるからである。

#### 2.1 論文の被引用数に及ぼす要因に関する研究

本節では、論文の被引用数に影響を及ぼす要因に関する多くの研究をレビューし、その中で、各要因を表すのにどのような測度が用いられたか、それらの測度が被引用数に与える影響がどの程度であったかを示す。後述するように（3.1.2、3.2.1 参照）、論文の被引用数とその分野、タイプ、言語に依存することは明らかであるので、このレビューではこれらの要因は対象としないが、多くの要因の影響を検討した研究にこれらの要因が含まれている場合は言及することがある。

2.1.1 では内在的要因、2.1.2 では外在的要因と被引用数の関係について個別的に論じ、2.1.3 で多くの潜在的要因の影響を総合的に見た研究（多くは重回帰分析を使用）について述べる。

##### 2.1.1 内在的要因による論文の被引用数への影響

1.2 の冒頭で述べたように、本研究は、外在的要因が論文の被引用数にどの程度影響を与えるかを明らかにすることが目的なので、内在的要因の影響は対象としない。しかし、内在的要因を被引用数の関係を論じた先行研究は少なくないので、ここではそれらについてレビューを行う。特に、内在的要因をどのような測度で表現しているかに注目する。

###### 2.1.1.1 論文の質に関する要因

専門家による評価が高い論文が高い引用を得るかどうかに関心した研究はかなりある。これらにおいては、次のようなピア評価指標が測度として用いられ、それらの被引用数への影響が調べられている：

## 第2章 先行研究のレビュー

- (a) 雑誌投稿論文の査読結果(Bornmann and Daniel, 2008; Patterson and Harris, 2009; Bornmann et al., 2011)
- (b) 謝辞に言及された競争的研究資金の獲得(Cronin and Shaw, 1999; Haslam et al., 2008; Rigby, 2013)
- (c) 著者の自己評価(Aksnes, 2006)。
- (d) 外科臨床研究の方法の厳密性に関する外部者の評価(Akcan et al., 2013)
- (e) 臨床的適合性とニュース価値に関する外部者の評価 (Callaham et al., 2002; Lokker et al., 2008)
- (f) EBM 誌への採録(Lokker et al., 2008)

これらの測度は、(b)を除けば論文から直接得られるものではない。なお、(d)、(e)、(f)は臨床医学論文を対象としたものである。

論文の質を自己評価やピア評価以外の定量的測度で表すことは難しいが、Chen (2012)による研究は、その点から注目される。彼は、論文の潜在力（または価値）はそれが現状の知識構造を改変する程度（boundary-spanning 能力）で表されるとし、この能力を、既存の知的ネットワーク構造の変化を示す次の3つの定量的測度で可操作化した：(a)モジュラリティ変化率、(b)クラスターリンケージ、(c)中心性拡がり。これら3つの「内在的」測度と、従来からよく使われる3つの「外在的」測度（著者数、参考文献数、ページ数）を説明変数として、いくつかの異なる分野の文献集合に対する負の2項重回帰により、被引用数を予測した。その結果、クラスターリンケージは3つの「外在的」変数よりずっと説明力の強い予測変数であること、中心性拡がりも予測力がやや不安定であるが、boundary-spanning 能力を持つことを示した。

### 2.1.1.2 論文が扱う方法や得られた結果に関する要因

仮説検定論文における検定方法、検定結果によって引用インパクトに差が生ずるか否かを論じたいくつかの研究がある。これらの研究では、次のような測度と被引用数の関係が検討されている：

- (a) サンプルサイズまたは被験者のタイプ(Callaham et al., 2002; Leimu and Koricheva, 2005; Lortie et al., 2013)
- (b) 対照群や無作為化の有無(Callaham et al., 2002)
- (c) 結果が肯定的か否定的か、あるいは仮説の支持/不支持(Callaham et al., 2002; Leimu and Koricheva, 2005; Fanelli, 2013)
- (d) 統計的有意性の強さ(Leimu and Koricheva, 2005)。

総じて、これらと被引用数の間に有意な関係は見出されないか、あっても弱いものであった。

### 2.1.1.3 論文が扱うトピックやテーマに関する要因



## 第2章 先行研究のレビュー

2.1.3 に挙げる多くの要因の影響を統合的に論じた研究のいくつかでは、論文の扱うトピックやテーマに関する属性が重回帰分析の説明変数に含まれている。それらの属性には、医学論文のトピック用語(Fu and Aliferis, 2010)、臨床医学論文の研究デザイン(Lokker et al., 2008)、地球科学論文の細分野や研究手法(Stewart, 1983)、人口学論文のテーマ(van Dalen and Henkens, 2001, 2005)等がある。

### 2.1.1.4 論文の学際性及び国際性に関する要因

論文の学際性や国際性の程度の可操作化も難しい問題であるが、参考文献を用いる方法が比較的によく使われる。Lariviere and Gingras (2010a)は、2000年のWeb of Scienceの全論文を14の分野に分け、論文の属する分野と異なる分野の参考文献の比率をその論文の学際性の測度として、学際性が被引用数に及ぼす効果を分析した。被引用数の学際性に対する依存性のパターンは分野により異なるが、どの分野でも、学際性が高いか低い両端の論文では被引用数が低くなる傾向があった。Didegah and Thelwall (2013)は、ナノ科学・ナノテクノロジー分野の論文への影響要因の分析において、参考文献の著者所属国の多様さを国際性の測度として、それが有意な予測変数の一つとなることを示した。

表 2-1 は、2.1.1 で述べた研究の要約である。

## 2.1.2 外在的要因による論文の被引用数への影響

### 2.1.2.1 共同研究は被引用数を押し上げるか?—研究協力に関する要因

共同研究は単独の研究に比べて研究の影響範囲が広まり、それだけ引用も増えると考えられることができる。更に、同じグループ内、機関内の共同研究より、機関間共同研究や国際共同研究が引用を誘発するとも考えられる。これらの仮説に基づき、論文の被引用数に及ぼす共同研究の効果を検討した研究はかなりの昇る。共同研究の規模や拡がりの程度を可操作化する測度としては、論文の著者数、著者が所属する異なる機関の数(所属機関数)、所属機関が属する異なる国の数(所属国数)がよく用いられる。また、論文を単独著者研究、同一機関内共同研究、同一国内の機関間共同研究、国際共同研究に分類する方法もある。

著者数、所属機関数、所属国数のすべてを考慮して被引用数の関係に注目した研究に、Basu and Lewison (2005)、Figg et al. (2006)、Sooryamoorthy (2009)があり、概ね、共同研究により被引用数が押し上げられるという結果が得られている。

## 第2章 先行研究のレビュー

表 2-1 論文の質や内容に関する要因の引用への影響を扱った研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測定 <sup>b)</sup>						
				ピア評価	自己評価	研究資金獲得	独創性指標	方法・結果	テーマ・トピック	学際性・国際性
Akcan et al. (2013)	臨床医学	192	Kruskal-Wallis 検定: 順位相関	C						
Aksnes (2006)	一般	1549	順位相関		B					
Bornmann & Daniel (2008)	化学	1,837	NBMR	A						
Bornmann et al. (2011)	天文学	315	カイ2乗検定	A						
Callaham et al. (2002)	救急医学	204	CMR	B				B		
Chen (2012)	4つのトピック	トピックにより 1,300 - 6,800	NBMR, ZINBMR				A			
Cronin & Shaw (1999)	図書館情報学	716	カイ2乗検定			C				
Didegah & Thelwall (2013)	ナノ科学・技術	50,162	ZINBMR							A
Fanelli (2013)	仮説検定研究	2545	NBMR					B		
Fu & Aliferis (2010)	内科学	3,788	SML, LogMR						B	
Haslam et al. (2008)	社会心理学・人格心理学	308	LMR (LogC)			B		C		
Larivière & Gingras (2010a)	一般	2000年の 全WoS論文	単純比較							B
Leimu & Koricheva (2005)	エコロジー	228	t検定; ANOVA; 相関					C		
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )	A				A		
Lortie et al. (2013)	エコロジー・進化生物学	1332	一般化線型モデル					C		
Patterson & Harris (2009)	生医学における物理	1095	相関	A						
Rigby (2013)	生化学	3596	LMR (Log(C+1))			B				
Stewart (1983)	地球科学	139	LMR (logC)						A	
Van Dalen & Henkens (2001)	人口学	1,371	NBMR						A	
Walters (2006)	犯罪心理学	428	NBMR							C

a) LMR:線形重回帰 CMR:カテゴリカル重回帰 NBMR:負の2項重回帰 ZINBMR:Zero-inflated 負の2項重回帰 LogMR:ロジスティック重回帰 SML:教師付き機械学習

LogC, Log(C+1),  $\sqrt{C}$  は LMR の従属変数を示す。

b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意なまたは負の影響

Persson et al. (2004) は、過去 20 年間の論文あたり参考文献数の増加（従って被引用数の増加につながる）の原因の一つが共著の増大傾向にあるという仮説を確かめるために、1980 年と 2000 年における論文の著者数と、これらの論文が得た被引用数の関係を調べた。両年とも、著者数が一人増えると平均被引用数が約 0.6 上昇する（発表後 3 年間の引用ウィンドウで）という明確な関係が見られた。しかし、同じ著者数の論文における被引用数は、1980 年から 2000 年の間で 8 増えているので、共同研究効果だけでこの期間の被引用数の増加を説明することはできない。

この他にも、論文の著者数と被引用数の間に正の相関があるという多くの報告がある (Aksnes, 2003b; Leim and Koricheva, 2005; Bornmann and Daniel, 2006; Davis et al., 2008; Lokker et al, 2008; Sin, 2011; Chen, 2012; Peng and Zhu, 2012; Fanelli, 2013; Rigby, 2013)。しかし、多数の説明変数を含めたいくつかの重回帰分析では、著者数の被引用数予測力は弱い (Peters and van Raan 1994; van Dalen and Henkens, 2001; Walters, 2006; Bornmann and Daniel, 2008) か有意でない (Stewart, 1983; Fu and Aliferis, 2010)。

Hsu and Huang (2011) は、論文数が多くインパクトファクター(JIF)も高い 8 つの雑誌

## 第2章 先行研究のレビュー

について、論文の著者数の増加とともに統計的には被引用数は上昇する傾向があるが、より著者数の多い論文がより高い引用数を得る確率はそれほど高くはなく、雑誌により 53～65%程度であることを示した。Levitt and Thelwall (2009b)は、高被引用著者のみからの論文集合の中で著者数と被引用数の間に明確な関係が見られないことから、混合論文集合において見られる論文の著者数と被引用数の間の正の相関は、多くの共著者を持つ著者ほどその論文が多く引用を得るという傾向を反映したものであらうと示唆している。

国際共同研究論文は、単独機関論文や国内機関間共同研究論文よりよく引用されると言われる (Katz and Hicks, 1997; van Raan, 1998; Persson et al., 2004; Sooryamoorthy, 2009; Sin, 2011; Peclin et al., 2012; Ibáñez et al., 2013; Bordons et al., 2013)。たとえば Katz and Hicks (1997)は、同一機関や国内別機関の一人の著者との共著は被引用数を平均 0.75 上昇させるのに対し、一人の外国著者との共著は平均 1.6 の被引用数上昇をもたらすとした。しかし He (2009)は、これらの研究で行っている国際共著、国内他機関との共著、機関内共著というグループ化では、国際共著グループの平均著者数が最も多くなるため、そのグループに正のバイアスがかかっている可能性を指摘した。そして、ニュージーランドの 65 人の生医学研究者の 1,860 論文に対し、国外共著者数、国内他機関共著者数、同一機関共著者数を手作業で同定し、負の 2 項重回帰分析を行った結果では、一人の国外共著者の増加と一人の同一機関共著者の増加が被引用数増にもたらす効果は同等であった (国内他機関共著者の効果は有意でない)。

表 2-2 は、ここで述べた研究の要約である。

### 2.1.2.2 著者、機関、国のハロー効果は存在するか？—著者に関する要因

高い名声や過去の業績を持つ著者、あるいは著名な研究機関に所属する著者による論文が、そうでない著者の論文よりも多くの引用を引きつけるというハロー効果の存在に関する多くの議論がある。

著者の名声、業績の可操作化は難しく、いろいろな測度が考えられる。それらには、当該論文発表前の論文数、それらの論文が得た引用に基づく指標、活動の期間、現在の職位などがある。

2.1.3 で述べる研究の多くで、これらが予測モデルの説明変数に使われている。過去の論文数については、Peters and van Raan (1994)、Haslam et al. (2008)が有効な予測変数としているが、Fu and Aliferis (2010)では有意な結果は得られていない。過去論文の引用に関しては、合計被引用数(Fu and Aliferis, 2010)、論文の平均被引用数(Stewart, 1983; Walters, 2006)、h 指数(He, 2009)、ISI Highly Cited list に載った著者の割合(Bornmann and Daniel, 2008)等が説明変数とされ、いずれも有意な予測変数という結果が出ている。シニアの著者による論文が高い引用数を得るという報告(Slyder et al., 2011; Peng and Zhu, 2012)と、その関係を否定する報告(Stewart, 1983; He, 2009)がある。

表2-2 研究協力に関する要因の引用への影響を扱った研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測度 <sup>b)</sup>		
				著者数	所属機関 数 <sup>c)</sup>	所属国数 <sup>d)</sup>
Aksnes (2003a)	自然科学	46,849	単純比較	A		A
Basu & Levison (2005)	天文学・天体物理学	95,186	LMR (Log(C+1))	A	A	C
Bordons et al. (2013)	薬学・薬理学	1,971 and 2,858 (2標本)	CMR	A	A	A
Bornmann & Daniel (2006)	生医学	1,586	NBMR	A		
Bornmann & Daniel (2008)	化学	1,837	NBMR	B		
Chen (2012)	4つのトピック	トピックにより1,300 - 6,800	NBMR, ZINBMR	A		
Davis et al. (2008)	生理学	1,619	NBMR	A		
Didegah & Thelwall (2013)	ナノ科学・技術	50,162	ZINBMR	C	A	C
Fanelli (2013)	仮説検定研究	2,545	NBMR	A		
Figg et al. (2006)	医学	164 - 886 (6標本)	LR (LogC)	A	A	
Fu & Aliferis (2010)	内科学	3,788	SML, LogMR	C	C	
Haslam et al. (2008)	社会心理学・人格心 理学	308	LMR (LogC)	C		
He (2009)	生医学	1,860	NBMR	A		A
Hsu & Huang (2011)	自然科学	10,000 - 15,000 (8標本)	LR	A		
Ibáñez et al. (2013)	計算科学	c.a. 20,000	Mann-Whitney検 定; Kruskal-Wallis 検定	C		A
Katz & Hicks (1997)	一般	c.a. 376,000	LR	A	A	A
Leimu & Koricheva (2005)	エコロジー	228	t検定; ANOVA; 相 関	A		
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )	A		
Peclin et al. (2012)	自然科学	5,263	ANOVA		B	A
Peng & Zhu (2012)	インターネット研究	18,580	LMR ( $\sqrt{C/y}$ )	A		
Persson et al. (2004)	一般	1980と2000の 全WoS論文	LR	A		A
Peters & van Raan (1994)	化学工学	226	CMR	B		
Rigby (2013)	生化学	3,596	LMR (Log(C+1))	A		C
Sin (2011)	図書館情報学	7,489	LogMR	A	C	A
Sooryamoorthy (2009)	一般	11,196	LR (LogC)	A	C	A
Stewart (1983)	地球科学	139	LMR (logC)	C		
Van Dalen & Henkens (2001)	人口学	1,371	NBMR	B		
van Raan (1998)	天文学	2,090	単純な平均比較		C	A
Walters (2006)	犯罪心理学	428	NBMR	B		

a) LR:線形単回帰 LMR:線形重回帰 CMR:カテゴリーカル重回帰 NBMR:負の2項重回帰 ZINBMR:Zero-inflated 負の2項重回帰 LogMR:ロジスティック重回帰 SML:教師付き機械学習

LogC, Log(C+1),  $\sqrt{C}$ ,  $\sqrt{C/y}$ はLMRの従属変数を示す。

b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響力

c) 国内機関共著の有無、他の国内機関の共著者数を含む。

d) 国際共著の有無、他国の共著者数を含む。

## 第2章 先行研究のレビュー

Danell (2011)は、エピソード記憶と Bose-Einstein 凝縮という 2 つの狭い領域の論文を対象に、著者の業績（過去の論文数とそれらの論文が得た引用数）から、それらの論文が発表後に得る引用数を予測できるかについて検討した。パーセンタイル回帰により、過去の論文数は、従属変数(将来の被引用数)のメディアン付近の少数のパーセンタイル値のみで有意な予測因子であるに過ぎないが、過去論文の被引用数はほとんどのパーセンタイル値で有意な予測因子であり、高いパーセンタイル値ほど有意性が強いことを見出した。

著者のステータスと論文の被引用数の関係を見た別の例として、Bornmann and Daniel (2006)がある。この研究では、ある財団のフェローシップに申請した若手研究者の論文が得た引用（申請の前後にわたる）を、申請合格者と不合格者の間で比較し、前者の被引用数が有意に多いことを示した。

著者が所属する機関のステータスを示す指標として、Leimu and Koricheva (2005)や Fu and Aliferis (2010)は Academic Ranking of World Universities (ARWU)のランクを、Stewart (1983)は組織の発表論文数を用いたが、これらの指標の被引用数への影響要因は見られないか、極めて弱いという結果である。

被引用数が特定の国に偏る傾向について、著者は自国の研究者による論文を引用しがちなため、米国等主要な論文生産国の著者の論文は高い引用を得る可能性が高いと言われる。Leimu and Koricheva (2005)は、エコロジーに関する仮説論文集合を用いて、英語圏著者論文が非英語圏著者論文に比して、また米国著者論文が欧州著者論文に比して年あたりの被引用数が有意に高いことを示した。Cronin and Shaw (1999)は、図書館情報学分野の論文で、第一著者が米・英・加の場合、それ以外の国の場合に比べて無被引用論文の割合が低いことを示した。これ以外にも、米国あるいは西・北欧の著者の論文がそれ以外の地域の著者の論文に比べてよく引用されることを示す多くの報告がある(van Dalen and Henkens, 2001, 2005; Basu and Lewison, 2005; Walters, 2006; Davis et al., 2008; Sin, 2011; Peng and Zhu, 2012)。しかし、著者所属国は被引用数に対する重要な予測変数ではないとする報告もある(Peters and van Raan, 1994; Haslam et al., 2008; Lokker et al., 2008)。

Pasterkamp et al. (2007)は、心臓学分野の 6 雑誌に 1996 年に発表された論文の責任著者の所属国と、その論文の参考文献の責任著者の所属国の関係を検討した。その結果、自己引用を除いても、自国の著者の論文の引用は期待値より 32%も過剰であった。自国引用バイアスは、どの国についても、どの雑誌にも存在した。Schubert and Glänzel (2006)も自国引用への偏りが見られると報告しているが、自国(または自言語)引用率の数え方を適正化すればそのような傾向は見られないとする研究もある(Bookstein and Yitzhaki, 1999; Moed 2005, pp. 291-300)。

表 2-3 は、ここで述べた研究の要約である。

表 2-3 著者に関する要因の引用への影響を扱った研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測度 <sup>b)</sup>				
				著者の生産性	著者の引用影響度	著者の経歴・活動期間	所属機関特性	所属国特性
Basu & Levison (2005)	天文学・天体物理学	95,186	LMR (Log(C+1))					A
Bornmann & Daniel (2006)	生医学	1,586	NBMR			A		
Bornmann & Daniel (2008)	化学	1,837	NBMR		A			
Cronin & Shaw (1999)	図書館情報学	716	カイ2乗検定					A
Danell (2011)	2つのトピック	728 and 1,450	パーセントイル回帰	B	A			
Davis et al. (2008)	生理学	1,619	NBMR					A
Fu & Aliferis (2010)	内科学	3,788	SML, LogMR	C	A		C	
Haslam et al. (2008)	社会心理学・人格心理学	308	LMR (LogC)	A			C	C
He (2009)	生医学	1,860	NBMR		B	C		
Leimu & Koricheva (2005)	エコロジー	228	t検定; ANOVA; 相関				B	A
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )					C
Peng & Zhu (2012)	インターネット研究	18,580	LMR ( $\sqrt{C/y}$ )			A		A
Peters & van Raan (1994)	化学工学	226	CMR	A				B
Sin (2011)	図書館情報学	7,489	LogMR					A
Slyder et al. (2011)	地理学・森林学	213	t検定; ANOVA; 相関			A		
Stewart (1983)	地球科学	139	LMR (logC)		A	C	B	
Van Dalen & Henkens (2001)	人口学	1,371	NBMR		A			B
Walters (2006)	犯罪心理学	428	NBMR		A		C	A

a) LMR:線形重回帰 CMR:カテゴリカル重回帰 NBMR:負の2項重回帰 LogMR:ロジスティック重回帰  
SML:教師付き機械学習

LogC, Log(C+1),  $\sqrt{C}$ ,  $\sqrt{C/y}$ は LMR の従属変数を示す。

b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響力

### 2.1.2.3 論文が雑誌のインパクトを決めるか、それとも逆か？—論文の発表誌に関する要因

論文の被引用数が、それが発表された雑誌のインパクトと正の相関を持つことは当然であるが、雑誌の名声やインパクトは質の高い重要な論文が高い引用を得た結果である (Seglen, 1994) のか、それともインパクト指標の高い雑誌の名声がその雑誌に載った論文に引用を引きつける (van Dalen and Henkens, 2005) のかについては、議論のあるところである。

2.1.3 でも述べるように、van Dalen and Henkens (2001, 2005) は、雑誌の名声の測度 (JIF、editorial board メンバーが得た引用インパクト、出版部数) が、論文の被引用数に大きな効果を持つことを示した。これ以外にも、2.1.3 で述べるいくつかの研究が、論文の引用を押し上げる最も重要な要因のひとつとして、それが発表された雑誌の JIF (あるいは他のインパクト指標) を挙げている (Peters and van Raan, 1994; Davis et al., 2008; Fu and Aliferis, 2010; Peng and Zhu, 2012; Bordons et al. 2013; Didegah

## 第2章 先行研究のレビュー

and Thelwall, 2013)。しかし Walters (2006)では JIF の説明力はやや弱かった。

Callaham et al. (2002)、Aksnes (2003b)、Bornmann and Daniel (2006)、Slyder et al. (2011)、Ibáñez et al. (2013) も、論文の被引用数とそれが発表された雑誌 JIF の間に相関があることを示した。更に、Lariviere and Gingras (2010b)は、論文タイトル、第一著者名、参考文献数が完全に一致するが別の雑誌に発表された「重複論文」4,532 対を用い、各対においてより高い JIF の雑誌に発表された論文とより低い JIF の雑誌に発表された論文を比較するというユニークな方法で、JIF の高い雑誌に発表された論文の方が平均して 2 倍の引用を得ることを示した。ほぼ同一の論文の間でこのように明白な差が見られたことは、雑誌の名声が論文の引用に及ぼすハロー効果が、少なくとも部分的に存在することを明らかにしたといえる。

表 2-4 は、ここで述べた研究の要約である。

表 2-4 雑誌に関する要因の引用への影響を扱った研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測度 <sup>b)</sup>			
				JIF	他のインパクト指標	出版部数、収録DB数等	学際性・国際性
Aksnes (2003a)	自然科学	46,849	単純比較	A			
Bordons et al. (2013)	薬学・薬理学	1,971 and 2,858 (2標本)	CMR		A		
Bornmann & Daniel (2006)	生医学	1,586	NBMR	A			
Callaham et al. (2002)	救急医学	204	CMR	A			
Davis et al. (2008)	生理学	1,619	NBMR	A			
Didegah & Thelwall (2013)	ナノ科学・技術	50,162	ZINBMR	A			B
Fu & Aliferis (2010)	内科学	3,788	SML, LogMR	A			
Ibáñez et al. (2013)	計算科学	c.a. 20,000	Mann-Whitney検定; Kruskal-Wallis検定	A			
Larivière & Gingras (2010b)	一般	4,532対	t検定	A			
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )			A	
Peng & Zhu (2012)	インターネット研究	18,580	LMR ( $\sqrt{C/y}$ )	A			B
Peters & van Raan (1994)	化学工学	226	CMR		A		
Slyder et al. (2011)	地理学・森林学	213	t検定; ANOVA; 相関	A			
Van Dalen & Henkens (2001)	人口学	1,371	NBMR	A		A	
Walters (2006)	犯罪心理学	428	NBMR		B		

a) LMR:線形重回帰 CMR:カテゴリーカル重回帰 NBMR:負の2項重回帰 ZINBMR:Zero-inflated 負の2項重回帰 LogMR:ロジスティック重回帰 SML:教師付き機械学習

$\sqrt{C}$ ,  $\sqrt{C/y}$ は LMR の従属変数を示す。

b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響力

2.1.2.4 参考文献に関する要因

2.1.3でも述べるように、多くの引用影響要因を考慮するいくつかの重回帰分析研究で、参考文献数は説明変数の一つに加えられ、有意な予測変数であることが示されている(Stewart, 1983; Peters and van Raan, 1994; Davis, et al., 2008; Haslam et al., 2008; Lokker et al. 2008; He, 2009; Didegah and Thelwall, 2013)。他の多くの研究でも、参考文献が多い論文の方がよく引用されるという結果が得られている(Chen, 2012; Rigby, 2013; Bordons et al., 2013)。

参考文献中に占める自己引用の比率や、参考文献の引用年齢や主題分野の分布等、参考文献のより細かい特徴も論文の被引用数に関係するのではないかと考えられるが、データを得るのに手間を要するためか、そのような研究は多くない。しかし、被引用数を予測するいくつかの重回帰分析研究で、潜在予測要因に参考文献の若さを含めている。Stewart (1983)は引用前3年以内、Peters and van Raan (1994)は引用前5年以内 (Price 指数)に発表された参考文献の比率が、いずれも比較的弱い予測変数であることを示した。Haslam et al. (2008)は、参考文献の平均年齢が短い論文の方が高い引用を得る傾向があることを示した。

表 2-5 参考文献に関する要因の引用への影響を扱った研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測度 <sup>b)</sup>		
				参考文献数	参考文献の 最新性	参考文献の インパクト
Bordons et al. (2013)	薬学・薬理学	1,971 and 2,858 (2標本)	CMR	A		
Chen (2012)	4つのトピック	トピックにより1,300 - 6,800	NBMR, ZINBMR	A		
Davis et al. (2008)	生理学	1,619	NBMR	A		
Didegah & Thelwall (2013)	ナノ科学・技術	50,162	ZINBMR	A		A
Haslam et al. (2008)	社会心理学・人格心理学	308	LMR (LogC)	A	A	
He (2009)	生医学	1,860	NBMR	A		
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )	A		
Peng & Zhu (2012)	インターネット研究	18,580	LMR ( $\sqrt{C/y}$ )			A
Peters & van Raan (1994)	化学工学	226	CMR	A	A	
Rigby (2013)	生化学	3596	LMR (Log(C+1))	A		
Stewart (1983)	地球科学	139	LMR (logC)	A	B	

a) LMR:線形重回帰 CMR:カテゴリカル重回帰 NBMR:負の2項重回帰 ZINBMR:Zero-inflated 負の2項重回帰  $\text{Log}C$ ,  $\text{Log}(C+1)$ ,  $\sqrt{C}$ ,  $\sqrt{C/y}$ は LMR の従属変数を示す。

b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響力



## 第2章 先行研究のレビュー

参考文献のその他の特性に関して、Peng and Zhu (2012)及び Didegah and Thelwall (2013)は、参考文献のインパクト(高被引用論文の比率あるいは平均被引用数)が高い論文が高い引用を得ることを示した。

表 2-5 は、ここで述べた研究の要約である。

### 2.1.2.5 論文の可視性に関する要因

これに関する測度として最もよく使われるのは、論文の長さである。単純にページ数が用いられることが多いが、雑誌により1ページあたり文字数は異なるので、その補正を行うことが望ましい。

論文の長さと同被引用数の間には、正の相関を報告するもの(Stewart, 1983; Peters and van Raan, 1994; Leim and Koricheva, 2005; van Dalen and Henkens, 2001, 2005; Haslam et al., 2008; Peng and Zhu, 2012)と有意な相関が見られないとするもの(Walters, 2006; Davis et al., 2008; Slyder et al., 2011; Rigby, 2013)がある。一方、Lokker et al. (2008), He (2009), Chen (2012)は逆に負の関係を得たが、これは重回帰分析に用いた他の説明変数が影響している可能性がある。

表 2-6 論文の可視性に関する要因の引用への影響を扱った研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測度 <sup>b)</sup>					
				論文長	図の数	表の数	タイトルの長さ	抄録の長さ	抄録の可読性
Bornmann & Daniel (2008)	化学	1,837	NBMR	C					
Chen (2012)	4つのトピック	トピックにより1,300 - 6,800	NBMR, ZINBMR	C					
Davis et al. (2008)	生理学	1,619	NBMR	C					
Gazni & Didegah (2011)	高被引用の5機関	124,937	相関						A
Haslam et al. (2008)	社会心理学・人格心理学	308	LMR (LogC)	A	C	C	B		
He (2009)	生医学	1,860	NBMR	C					
Leimu & Koricheva (2005)	エコロジー	228	t検定; ANOVA; 相関	A					
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )	C				C	
Peng & Zhu (2012)	インターネット研究	18,580	LMR ( $\sqrt{C/y}$ )	A					
Peters & van Raan (1994)	化学工学	226	CMR	B					
Rigby (2013)	生化学	3596	LMR (Log(C+1))	C					
Snizek et al. (1991)	3つのトピック	トピックにより8 - 20	LMR	C	C	C		C	B
Stewart (1983)	地球科学	139	LMR (logC)	A					
Van Dalen & Henkens (2001)	人口学	1,371	NBMR	A					
Walters (2006)	犯罪心理学	428	NBMR	C					

a) LMR:線形重回帰 CMR:カテゴリカル重回帰 NBMR:負の2項重回帰 ZINBMR:Zero-inflated 負の2項重回帰 LogC, Log(C+1),  $\sqrt{C}$ ,  $\sqrt{C/y}$ はLMRの従属変数を示す。

b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響力

## 第2章 先行研究のレビュー

論文の長さ以外では、図や表の数、タイトルや抄録の特徴（長さ、可読性等）と被引用数の関係を調べたものがある。Snizek et al. (1991)は、図表の数、著者抄録の語数や可読性と被引用数の関係について調べているが、いずれも被引用数との明白な関係は認められなかった。しかし、Gazni and Didegah (2011)は、Flesch の抄録可読性が低い（読みにくい）抄録の論文の被引用数が一般に高いことを示した。Haslam et al. (2008)による重回帰分析では、多くの説明変数の中に図表の数やいくつかのタイトルの特性を含めているが、その中ではタイトルの長さだけが被引用数と若干の関係があった。

表 2-6 は、ここで述べた研究の要約である。

### 2.1.3 多くの潜在的影響要因を総合的に考慮した引用分析

論文の被引用数に影響を与え得る要因として様々な特性が調べられてきたが、それらの研究の多くは、ある単一の要因に着目しているか、複数の要因をそれぞれ独立に見ている。従って、ある要因が被引用数と相関があっても、それが他の要因の交絡によるという可能性を否定できない。たとえば、著者数、著者所属機関数、論文の長さ、参考文献数などは、被引用数と正の相関があると多くの研究で報告されている。しかし、そもそもこれらの要因間に正の相関があるとすれば、他の要因の寄与を分離してそれぞれの要因を評価してみると、ある一部の要因のみが被引用数と真の有意な相関を持つ可能性が大きいことに注意すべきである。

重回帰分析は、個々の因子（独立変数）による効果を分離し、被引用数に対し高い説明力を持つ因子を推察するのに最もよく使われる方法である。この細節では、この方法を用いたいくつかの主な研究の概要を示す。これらの研究は、既に 2.1.2 の各細節でそれぞれ注目した要因に関して言及しているが、いろいろな要因を総合的に考慮することにより交絡の可能性を低め、結果の信頼性は比較的高いと考えられるので、ここで総合的観点から概説する。

#### 2.1.3.1 論文の被引用数に影響を及ぼす主要な要因を検討する研究

Peters and van Raan (1994)は、論文の被引用数に影響する要因についての研究は数多く存在するが、多くの要因の中から主要な決定要因が何であるかを特定しようとする研究は皆無に近いことを指摘し、化学工学分野を選んで、種々の要因がどの程度論文の被引用数に影響するかについて検討した。この分野の多産で国際的な名声もある 18 人の研究者を選び、これらの研究者が 1980–1982 年に発表した論文 ( $n = 226$ ) のそれぞれが、発表後 5 年以内に引用された回数を測定した。14 の要因を説明変数とする重回帰分析の結果、高度に有意な説明変数は次の 6 つであった（偏相関係数の高い順）：(a)著者のランク（1980–82 の論文数で測定）、(b)参考文献数、(c)言語、(d)発表雑誌の名声、(e)influence weight（雑誌影響度に関する Narin の指標）、(f)Price 指数。また、次の 4 変数も有意であった：

(g)記事主題分類(CA Section)、(h)ページ数、(i)著者数、(j)著者所属国。面白いことに、抜群に説明力の高い著者ランクは、被引用数との単純な相関では有意でなかった。このことは、影響を持ちそうな種々の要因を総合的に分析することの重要性を示すものである。

Didegah and Thelwall (2013)は、2007-2009の3年間に発表されたナノ科学・ナノテクノロジー分野の50,000件以上の論文を用いて、主要な被引用数決定要因を検討した。これまで考慮されていなかった要因として、雑誌の国際性と参考文献の国際性を含め（どちらも著者の所属国分布のGini係数を尺度とする）、8つの独立変数を選んだ。発表年ごと及び全期間の4つの文献集合に対して、Zero-inflated 負の2項(ZINB)回帰分析を行った。JIFと参考文献のインパクト（参考文献の平均被引用数で測定）が最も強い影響要因であり、この他、参考文献数、参考文献の国際性、所属機関数などの集合においても有意な予測要因であった。著者数の影響はほとんど見られなかった。著者所属国数と雑誌の国際性は負の効果を示す傾向があったが、これは、この分野では米国の研究が主導的であることと関係していると考えられる。

### 2.1.3.2 著者、雑誌、論文に関するどの要因が支配的か?

論文の被引用数に影響を与える潜在的要因を、著者の属性、雑誌の属性、論文自体の属性に分け、どの属性が支配的かを検討する研究がいくつか見られる。

最も早い時期ではStewart (1983)がある。1968年発表の地球科学論文139件が得た被引用数を、多数の著者属性と論文属性から重回帰分析により予測を試みた。著者属性のみによる回帰、論文属性のみによる回帰、及び全属性による回帰の決定係数( $R^2$ )の比較により、論文属性の方が著者属性より重要な影響要因であるとした。被引用数を増やす重要な要因は、論文属性では、参考文献数、論文の長さ、受理から出版までの早さ、参考文献の新しさ（やや弱い）であった。この他、論文の主題やタイプを表すいくつかのダミー変数（プレートテクトニクスに関するもの等）も被引用数に有意な影響を示した。一方、有意な著者属性は、全著者の過去の論文の平均被引用数、大学所属著者比率であった。著者数は有意な予測変数ではなかった。

Walters (2006)は、犯罪心理学分野の12誌に2003年に発表された428論文が得た被引用数を、著者属性、論文属性、雑誌属性を含む9つの説明変数を用いた負の2項重回帰により予測した。その結果、第一著者の過去論文の平均被引用数、第一著者が米国著者であること、レビュー論文であることが有意な予測変数であった。著者数（単著か共著かの2値）と雑誌のインパクト指標は有意に近く、第一著者の性や所属機関、論文の長さ、論文の主題（矯正心理学か犯罪心理学か）は有意でなかった。Waltersはこれらの結果から、著者属性が論文属性や雑誌属性より被引用数予測に重要であろうと示唆した。

Haslam et al. (2008)は、社会心理学・人格心理学の主要3誌に1996年に発表された308論文の被引用数を分析した。30の潜在要因を著者属性(author characteristics)、帰属

## 第2章 先行研究のレビュー

属性(institutional characteristics)、論文構成属性(article organization characteristics)、研究アプローチ属性(research approach characteristics)の4つに分け、まずそれぞれの属性内で重回帰分析を行い、次に、各回帰で有意となった9変数を用いて最終的な回帰を行った。その結果、(a)第一著者の生産性(過去の論文数)が高いこと、(b)第一著者より更に生産性の高い共著者が存在すること、(c)著名な雑誌に発表されること、(d)論文が長いこと、(e)参考文献が多いこと、(f)参考文献の平均年齢が若いことが、被引用数を押し上げる主な要因であるとした。最終回帰に選ばれたが有意でなかった変数には、第一著者以外の著者の総生産性、競争的研究資金の獲得、タイトルの長さ、レビューまたは理論論文であることなどがあった。

Peng and Zhu (2012)は、インターネット研究に関する18,580件の社会科学論文について、まず論文属性変数(著者属性を含む)により、次いで雑誌属性変数を加えて被引用数の重回帰予測を行った。その結果、雑誌属性、特にJIFの説明力が強いことを示した。しかし、論文属性でも、ページ数、著者数、トピックの関心度(抄録中のインターネット関係語数で測定)、参考文献中の高被引用論文比率、第一著者の活動年数が有意な予測変数であった。

### 2.1.3.3 論文への注目を惹く「シグナル」の研究

van Dalen and Henkens (2001, 2005)は、人口学分野の論文の被引用数に及ぼす要因を検討したが、その意図は、自然科学分野において引用インパクトに(引いては科学コミュニケーションに)影響する要因が社会科学である人口学にも適用されるかを明らかにすることであった。彼らは特に、論文への早期の注目を惹く「シグナル」としての、著者の名声及び雑誌の名声の役割に注目した。このため、この分野の1371論文のそれぞれが、発表後5年間(van Dalen and Henkens, 2001)及び10年間(van Dalen and Henkens, 2005)に引用された回数を数え、論文の著者、可視性、内容、掲載雑誌に関する様々な特性を説明変数とするいくつかの負の2項重回帰モデルを開発した。

JIF、出版部数、editorial boardsの名声(そのメンバーが得た平均被引用数により測定)のような雑誌の名声を示す変数は最も影響力の強い要因であった。それに対し著者の名声指標である著者(共著論文の場合は過去の引用が最多の著者)が過去に得た累積被引用数の影響力は、有意ではあるがそれほど強くなかった。高度に有意な相関を示した他の変数には、論文タイプ(短報やコメントは原著論文に比べ引用されない)、ページ数、論文が扱う地域(米国や欧州に注目した論文はよく引用される)、言語(仏語を用いる雑誌は英語のみの雑誌に比べ引用されない)がある。この他、著者の所属国、著者数、雑誌号中の掲載位置も中程度の有意な関係を示した。

### 2.1.3.4 その他の研究

上述の他にも、論文の被引用数に影響する可能性を持つ要因を検討した研究がいくつか

## 第2章 先行研究のレビュー

ある。これらは、ある特定の要因の被引用数に対する効果に着目して、その検討のため多くの制御変数を加えたものである。

Bornmann and Daniel (2008)は、査読制雑誌のピアレビューシステムが優れた論文を選択するという目的を果たしているか否かを検証するため、2000年に *Angewandte Chemie International Edition (ACIE)* に投稿された論文のうち、受理された878論文と拒絶され別誌に掲載された959論文の被引用数を比較した。他の種々の引用影響因子の影響を制御した負の2項重回帰分析の結果、受理論文は拒絶論文に対して40-50%平均引用数を上げるという結果を得た。制御要因の中では、言語（英語かそれ以外か）、著者のステータス（ISI Highly Cited.com にリストされた著者数）が統計的に有意であった。また、有機化学、物理・無機化学、高分子化学の論文は応用化学、生化学の論文より有意に被引用数が高かった（主題分類は CA Sections による）。著者数は、短期の引用ウィンドウの場合（発表後3年間）のみ有意であった。論文の長さは有意ではなかったが、これはここで対象とした論文が ACIE の Communications（一般に短い論文からなる）であったためであろう。

Lokker et al. (2008)は、EBM (evidence-based medicine)の観点からの有用性評価基準を通過した臨床医学論文の被引用数を、論文発表後3週間以内に得られる情報から予測できるかどうかを検討した。重回帰分析の説明変数には、通常の記事的特性の他、論文に対する質の評価に関するもの、雑誌の評価に関するもの等を含む。被引用数への有意な予測要因は、論文の質に関しては、EBM との適合性に関する専門家の評点（ニュース価値についての評点は有意性なし）、EBM 誌への抄録採録、雑誌評価に関しては、索引されるデータベース数、EBM の synoptic 誌への採択論文の割合、記事的特性については、著者数、参考文献数、多施設共同研究であること、治療論文であることであった。

Davis et al. (2008)は、オープンアクセス(OA)論文がそうでない(NOA)論文に比べてよく引用されるという多くの経験的結果が、著者の自己選択等 OA 化以外の効果によるものではないかという仮説の下に、米国心理学会が発行する雑誌の論文を半年の間 OA 論文と NOA 論文に無作為に振り分け、両者の間で被引用数を比較する無作為化比較試験(RCT)実験を行った。多くの制御変数を含む負の2項重回帰分析の結果、OA 論文の被引用優位性を示す証拠は見られなかった。用いた制御変数のうち著者数、米国著者の存在、JIF、参考文献数が有意、論文長は有意ではなかった。

He (2009)は、国際共著論文の引用インパクトが高いという通説に反論するため、ニュージーランドの生医学研究者の論文に対し、いくつかの共著者変数と制御変数を用いた負の2項重回帰分析を行った。その結果、同一機関の共著者の増加は海外（ニュージーランド以外の国）の共著者の増加と変わらない被引用数増大効果を持つことを示した。用いた制御変数のうち、参考文献数が多いこと、著者の年齢が低いことは有意に被引用数を押し上げるが、著者の h 指数はモデルにより有意または非有意、論文長と論文タイプは有意でないことを示した。

## 第2章 先行研究のレビュー

Fu and Aliferis (2010)は長期（発表後10年間）の被引用数を予測するため、SVM アルゴリズムを用いた教師付き機械学習モデルを開発した。医学総合誌8誌から抽出した1991-94年発表の内科学論文3,788件にこのモデルを適用して、多数の主題特性変数（各論文のタイトル、抄録、MeSH 索引語から抽出した用語）と9つの書誌変数から有効な変数を選択し、更にもの中から、高被引用論文と低被引用論文を識別するロジスティック回帰分析により有意な変数を決定した。書誌変数のうち、最終的に有効な予測変数とされたのはJIFと最終著者の累積過去被引用数のみで、論文タイプと第一著者の累積過去被引用数は、ロジスティック回帰で有意とならなかった。有効な予測変数とされた主題属性変数は、ロジスティック回帰での被引用数閾値によって大きく変化した。

2.1.3 で述べた研究を表 2-7 に要約する（この表の Chen (2012)については 2.1.1.1 で概説した）。

第2章 先行研究のレビュー

表 2-7 種々の引用影響要因を総合的に考慮した研究の概要

文献	対象分野	標本サイズ (n)	分析の方法 <sup>a)</sup>	要因を表す測定 <sup>b)</sup>																		
				研究の質	方法・結 果	論文の内 容	著者数	所属機関 数 <sup>c)</sup>	所属国数 <sup>d)</sup>	著者の生 産性	著者の引 用影響度	著者の経 歴・活動 期間	所属機関 特性	所属国特 性	雑誌のイ ンパクト	他の雑誌 属性	参考文献 数	参考文献 の最新性	参考文献 のインパ クト	論文長	図・表の 数	タイトル・ 抄録の特 性
Bornmann & Daniel (2008)	化学	1,837	NBMR	A			B														C	
Chen (2012)	4つのトピック	トピックにより 1,300 - 6,800	NBMR, ZINBMR	A			A								A						C	
Davis et al. (2008)	生理学	1,619	NBMR				A						A	A	A						C	
Didegah & Thelwall (2013)	ナノ科学・技術	50,162	ZINBMR			A	C	A	C					A	B	A						
Fu & Aliferis (2010)	内科学	3,788	SML, LogMR			B	C	C		C	A		C	A								
Haslam et al. (2008)	社会心理学・人格心理学	308	LMR (LogC)	B	C		C			A			C	C		A	A			A	C	B
He (2009)	生医学	1,860	NBMR				A				B	C			A						C	
Lokker et al. (2008)	臨床医学	1,261	LMR ( $\sqrt{C}$ )	A	A		A							C	A	A					C	C
Peng & Zhu (2012)	インターネット研究	18,580	LMR ( $\sqrt{C/y}$ )				A						A	A	A	B			A		A	
Peters & van Raan (1994)	化学工学	226	CMR				B			A				B	A		A	A			B	
Stewart (1983)	地球科学	139	LMR (logC)			A	C				A	C	B			A	B			A		
Van Dalen & Henkens (2001)	人口学	1,371	NBMR			A	B				A			B	A	A					A	
Walters (2006)	犯罪心理学	428	NBMR			C	B						C	A	B						C	

a) LMR:線形重回帰 CMR:カテゴリーカル重回帰 NBMR:負の2項重回帰 ZINBMR:Zero-inflated 負の2項重回帰 LogMR:ロジスティック重回帰 SML:教師付き機械学習

LogC,  $\sqrt{C}$ ,  $\sqrt{C/y}$ は LMR の従属変数を示す。

- b) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響
- c) 国内機関共著の有無、他の国内機関の共著者数を含む。
- d) 国際共著の有無、他国の共著者数を含む。

### 2.1.4 引用影響要因とそれらを表す測度のまとめ

以上の先行研究レビューに基づき、引用影響要因とそれらを表す主な測度、及びそれらの測度が被引用数に及ぼす影響の程度を、表 2-8 にまとめた。雑誌インパクト指標（JIF など）と参考文献数は強い影響要因であることが多くの研究で示されており、研究例はやや少ないが、著者の引用影響度も強く影響することが示唆されている。著者数と論文長も多くの研究で取り上げられているが、前者の影響度は中程度であり、後者の影響に対する支持は少ない。

表 2-8 引用影響要因とそれらを表す測度（先行研究のまとめ）

要因	測度	先行研究での検討 <sup>a)</sup>	被引用数との関係 <sup>b)</sup>		
			先行研究全般 <sup>c)</sup>	総合的分析研究 <sup>d)</sup>	
内在的 要因	質・内容要因	研究の質	B	C	A
		方法・結果	C	D	
		研究の内容	C	C	B
外在的 要因	研究協力要 因	著者数	A	B	C
		所属機関数	B	C	
		所属国数	A	B	
	著者要因	著者の生産性	D		
		著者の引用影響度	C	A	A
		著者の経歴・活動期間	C	C	
		所属機関特性	C	D	D
	雑誌要因	所属国特性	B	B	C
		雑誌インパクト指標	A	A	A
		他の雑誌特性	D		B
	参考文献要 因	参考文献数	B	A	A
		参考文献の最新性	D		
		参考文献のインパクト	D		
	論文可視性 要因	論文長	A	D	D
		図表の数	D		
タイトル・抄録の特性		D			

a) A：非常に多い B：かなり多い C：多くない D：僅少

b) A：非常に強い B：強い C：ある程度あり D：弱い

c) 「先行研究での検討」が D の測度は未記入

d) 総合的分析研究における事例が少ない測度は未記入



## 2.2 論文の引用履歴あるいは引用持続性に関する研究

### 2.2.1 論文の老化に関する数学モデルの開発

引用数の経時変化を表す数式、すなわち老化関数(*obsolescence function*)の研究は古くからのテーマであった。最も単純なのは指数型減衰への当てはめであるが、論文発表直後の引用は少なく、発表から数年後にピークに達するというのが一般的傾向なので、これは定性的にも実態をよく表さない。

Avramescu (1979)は、ある  $t$  で最大値をとる 2 種の老化関数  $c(t)$  (論文発表後の経過時間  $t$  における被引用数) を提案し、式中の 3 つのパラメータを調整することにより、多くの論文の引用履歴をある程度近似することができることを示した。Egghe and Ravichandra Rao (1992)は、老化関数の増減度を表す老化因子  $a(t) = c(t+1)/c(t)$  を考察し、実際の例では  $a(t)$  がある時間  $t$  で最小値をとることから、 $c(t)$  として対数正規モデルが適切であるとした。Burrell (2003)は、信頼性理論における失敗率関数を援用して被引用年齢分布を論じ、Egghe and Ravichandra Rao (1992)の結論が妥当であることを指摘した。

Glänzel and Schoepflin (1995)は、論文集合 (たとえばある年にある雑誌に発表された論文) の引用獲得プロセスに対する確率モデルから、論文が早期に受容される速度とある時間経過の後減衰する速度を示す指標を導いた。そして、減衰のパターンは個々の雑誌より分野に固有であること、減衰の遅いことは必ずしも受容が遅いこととは結びつかないことを示した。

Della Briotta Parolo et al. (2015)は 1960–1990 年の論文の発表年による引用減衰の変化を調べた結果、近年になるにつれて発表から引用のピークまでの間隔が縮まり、引用半減期が短縮していることを示した。また、ピーク年以降の引用減衰は逆冪型より指数型によくマッチした。

### 2.2.2 高被引用論文(HCP)とそれ以外の論文の引用持続性比較

Line (1984)の初期の研究により、高被引用論文は引用の持続性も高いことが示され、この研究が契機になって、高被引用論文(HCP)の被引用年齢分布が注目されるようになった。

Aversa (1985)、Cano and Lind (1991)は、彼らを選んだ HCP (Aversa は 1972 年に発表された 400 の HCP、Cano and Lind は Garfield が "Citation Classics" として選んだ 2 分野の HCP (Garfield 1984a, 1984b)) の長期間にわたる引用老化パターンを分析し、いずれも、短期引用型と長期引用型の 2 つのグループに分かれることを示した。Aversa は、

長期引用型の方が短期引用型より全期間被引用数が多いことも見出した。Aksnes (2003b) は、ノルウェー著者により 1981-89 年に発表された論文の中で、HCP はそうでない論文に比べ引用寿命が幾分長いことを示した。

Levitt and Thelwall (2008) は、HCP の被引用経年変化に関する論点について、より体系的な検討を行った。まず、1974 年に発表された HCP (自然科学 53 論文と社会科学 33 論文) の被引用経年変化が極めて多様で、Aversa (1985) や Cano and Lind (1991) が述べたような 2 つのパターンに分類できるものではないことを示した。次に、6 つの分野から 1969-1971 年発表の HCP を各 36 件採り、それらの 36 年間(1971-2006)の被引用年齢分布の分析から次のことを見出した: (a) どの分野でも、全期間累積被引用数に対する早期引用(最初の 6 年間)の比率は論文により大きく変動する; (b) 全期間被引用数が多い論文は早期引用率が低い傾向があるが、分野により相関の程度は異なる; (c) どの分野でも、HCP の引用持続性は非 HCP に比べ高い。更に彼らは、図書館情報学分野の HCP においても、より引用ランクの高い論文の方が遅延引用される傾向があることを見出した(Levitt and Thelwall, 2009a)。

### 2.2.3 引用ウィンドウ期間による論文の引用ランクの変化

Abramo et al. (2011a) は、2001 年にイタリアの著者によって発表された論文の 2008 年までの引用履歴を観測し、引用ウィンドウ期間を短くとしたときのインパクト測定の誤差を 8 つの分野ごとに検討した。たとえば、発表年を含む 3 年間の引用ウィンドウでの累積引用数は全期間 (9 年間の引用ウィンドウ) のその 65~88% (分野による) であり、3 年間ウィンドウと 9 年間ウィンドウの間の論文被引用数ランクの Spearman 順位相関係数は 0.79~0.96 であった。

Wang (2013) は、論文の短期の引用インパクトと長期のそれの間の相関を研究する中で、論文の引用獲得の早さを示す測度として Citation Speed を導入した。ある論文の発表された年を 0 とし、年 0 から年  $t$  までにこの論文が得た累積引用数を  $C(t)$  とすると、年  $T$  における Citation Speed は次式で与えられる。

$$\text{Citation Speed} = \frac{1}{T} \sum_{t=0}^{T-1} C(t) / C(T)$$

Citation Speed は [0,1] の値を取り、早期引用ほど値が 1 に近づく。1980 年に発表された論文の 31 年間の引用データから、論文の Citation Speed の分布は分野間のみならず同一分野内、更に同一雑誌内でも不均等であり、従って短期と長期の引用ウィンドウでの引用ランクの間にはかなりの不一致があった。また、最終被引用数が高くなるにつれて引用が遅延的になる傾向があった。

Abramo et al. と Wang はいずれも、発表後短期間の引用データから論文のインパクト評

価を行うことの問題を指摘し、研究評価を行う際の引用ウィンドウは、正確さとタイムリーさの間のトレードオフを考慮して慎重に選択すべきと主張している。

2.2.2 で言及した Levitt and Thelwall (2008)は、彼らが調査した6分野中5分野で、最初の6年間と全期間(36年間)のHCP被引用数ランクの間に有意な相関は見られないと報告している。

### 2.2.4 引用持続性による論文の分類及び引用持続性を測る指標

何人かの研究者が、引用持続性により論文を分類する試みを行っている。特に、長い無被引用期間を経た後注目される論文は、*delayed recognition papers* (Garfield, 1980; Glänzel et al., 2003)あるいは*sleeping beauties* (van Raan, 2004)と呼ばれ、興味が持たれている。逆に、発表直後によく引用されるがその後忘れられる論文に対する*flashes in the pan* (van Dalen and Henkens, 2005)という呼称もある。当初*flashes in the pan*と思われていたがある時点になって急に引用されるようになった論文に、*all-elements sleeping beauties* という名が与えられた (Li and Ye, 2012; Li, 2014)。

Glänzel et al. (2003)は、1980年に発表された論文の21年間の引用履歴から、60件の論文を遅延認識論文(*delayed recognition papers*)とした。一方 van Raan (2004)は、(a)睡眠期の年あたり被引用数、(b)睡眠期間、(c)覚醒後の年平均被引用数、から"*sleeping beauties*"の条件を定め、1980年以降の論文から該当する論文約360件を抽出した。どちらの場合も遅延認識論文は稀少であるが、Burrell (2005)は、彼の引用老化モデルに基づく計算の結果、"*sleeping beauties*"は確率的ランダム事象よりは多く出現し、それなりの理由が存在するとした。

Li et al. (2014)は、*sleeping* 期間中の引用年齢分布(彼らはこれを"*heartbeat spectrum*"と名付けた)に着目し、"*early heartbeats*"論文に比べ"*late heartbeats*"論文はずっと覚醒され易い(*sleeping beauty* になり易い)ことを見出した。

Lachance and Lariviere (2014)は、1963–1975年発表の論文のうち、発表後10年間以上全く引用がなくその後引用があった論文を"*sleeper*"とし、これに該当する論文(全論文の約5%)と同じ期間のそれ以外の論文(対照群)の間で引用パターンを比較した。睡眠終了後の*sleepers*の被引用年齢分布は、全体的には対照グループと似て徐々に減衰するが、高い引用を得る*sleepers*(すなわち*sleeping beauties*)では減衰は見られないか逆に上昇する傾向があり、対照群と異なる挙動を示すことを見出した。しかし、*sleeping beauties*はごく少数である上、それらの被引用数は対照群の高被引用論文のそれに比べずっと少ないので、この結論の一般化には注意が必要としている。

Li and Shi (2016)は、*sleeping beauties*に関する van Raan の基準(2004)は現在では適用できないとして、より長期にわたる引用履歴に基づく2つの新しい基準を示した。50年以上の引用履歴データを持つノーベル賞受賞者の論文からこれらの基準に合う25論文を

抽出したが、このうち van Raan の基準に合致するのは 10 件のみであった。

Baumgartner and Leydesdorff (2014)は、ノンパラメトリック統計手法の group-based trajectory modeling (GBTM) を適用して、論文集合を被引用年齢分布パターンの異なるいくつかの部分集合に区分する方法を提案した。この方法を 1996 年発表のいくつかの論文集合に適用し (いずれも 16 年間の引用履歴を使用)、それぞれを 3~7 のグループに分けることができた。これらのグループは、全被引用数の高、中、低だけでなく、引用が早期的还是持続的かによっても区分される。但し、グループ数の決定に主観的判断が必要、モデルを単純化するため外れ値の除去や初期値の設定が必要等の限界があり、多くの雑誌を含んだ大規模な論文集合への自動的適用は難しいとしている。

Wang et al. (2015)は、1 から Citation Speed (2.2.3 参照) を引いたものを Citation Delay と定義し、2001 年に発表された論文の長期 (13 年間) 引用データに対して、この測度が論文の学際性とそれ以外のいくつかの論文属性にどのように影響されるかを分析した。この Citation Delay は、引用年齢分布の全域の情報を用いた定量的な持続性指標として、これまでに提案された唯一のものである。なお、Li et al. (2014)が導入した  $G_s$  も Citation Delay とほぼ同じ考えに立っているが、彼らはそれを heartbeat spectrum (上記参照) の測定にのみ用い、一般的な引用持続性指標とはしていない。

これとは別に、Ke et al. (2015)は睡眠の程度を定量的に測る指標 "Beauty Coefficient" を導入した。この指標は、引用がピークを迎える年での被引用数とそれまでの毎年の被引用数との格差を測るものである。2 つのデータコレクションにおける長期的引用履歴データから得られたこの指標の分布は、どちらも逆冪型分布であった。このことから彼らは、sleeping beauty を識別する明確な境界を示すことはできないと結論した。

### 2.2.5 引用持続性と論文の他の性質との関係

長期的被引用数以外の論文特性と引用持続性との関係についての研究は少なく、van Dalen and Henkens (2005)、Costas et al. (2010)、Wang et al. (2015)を見る程度である。van Dalen and Henkens (2005)と Costas et al. (2010)はいずれも、論文を引用僅少型 [I]、早期引用型 (flash in the pan) [II]、遅延引用型 (sleeping beauty または delayed) [III]、標準型 (normal) [IV] の 4 つのクラスに分け、それらのクラスに属する論文の特性を比較した。また、Wang et al. (2015)は、彼らが提案した引用持続性測度 Citation Delay に論文の学際性その他の特性が及ぼす影響を検討した。しかし、van Dalen and Henkens (2005)ではクラス II、III、IV のそれぞれとクラス I との比較であり、3 クラス間の違いは明示的には報告されていない。また、Costas et al. (2010)と Wang et al. (2015)は、研究の主要な目的が別にある (Costas et al. は持続性による一般的な論文分類の方法論の開発、Wang et al. は論文の引用持続性と学際性の関係の検討)、論文の種々の属性と引用持続性の関係は比較的簡単にしか述べられていない。

これらの研究の結果と本章での研究の結果は、第5章で詳しく比較・考察する(5.4.2参照)。

### 2.3 論文の著者識別方法に関する研究

第3章第3節で論ずる著者の過去論文数の測定において肝要なことは、著者名サーチで検索された論文の著者が、意図している著者か同名の別人かを判別することである。ここでは、この著者識別方法に関する研究をレビューする。

著者識別の最良の方法は、検索対象の研究者の発表論文リストと照合することである(Rinia et al., 1998; van Raan, 2006; Bornmann and Daniel, 2007)が、この方法は、限られた機関の少数の研究者以外に適用することは難しい。また、非常に多数の著者について分析を行う場合には多大な労力を要する。よく用いられるのは、データベースに記載されている著者所属機関や論文の主題から判断する方法である。しかしこれも、大量の論文に対して人手で行うことは非現実的である。また、これらの方法で必ずしも正しい答が得られる訳ではない。同名の研究者が同じ機関に所属することや同じテーマの研究を行うことはあるし、その一方で、同じ研究者が所属やテーマを変えることもある。

同名の著者の識別に、共著者の情報が極めて有効であることが、いくつかの研究から明らかにされている。Wooding et al. (2006)は、論文の研究テーマと助成機関の情報に、著者の共著関係の情報を組み合わせたアルゴリズム的方法により、再現率(対象著者の論文中正解とされた比率)99%、適合率(正解とされた論文中対象著者論文の比率)97%を達成したと報告している。彼らの対象著者は、特定の1機関の助成を受けた狭い専門領域の研究者であるために、このような高い正解率が達成できたのかもしれない。

Kang et al. (2009)は、情報技術関係の会議録論文に現れる韓国の同名著者を、彼らの共著者に基づいてクラスタリングする方法を検討した。彼らは、対象とした論文に現れる「明示的」(explicit)共著者にWebから得られる暗示的(implicit)共著者を加えることにより情報を強化し、再現率87%、適合率88%の著者識別を達成した。

いくつかの研究グループは、共通の著者名(多くの場合姓とfirst nameイニシアル)を含む書誌レコードを個々の異なる著者のレコードにクラスター分割するための識別法を開発している。これらの方法は、書誌レコードにどんなメタデータ項目を含むか、あるいはWeb等の外部情報源からの情報を加えるか否かによりそれぞれ異なる。

Gilesのグループは、2つの教師付き方法-naïve Bayesモデルとsupport vector machine (SVM)モデル(Han et al., 2004)と3つの教師なし方法-K-way spectral clusteringモデル(Han et al., 2005)、SVM-DBSCANモデル(Huang et al., 2006)、topic-basedモデル(Song et al., 2007)-を提案した。最後の2つの方法では、「推移性の破れ」

(transitivity violation)<sup>10</sup>の問題を考慮に入れた。彼らは、DBLP (Digital Bibliography & Library Project)と CiteSeer から抽出した同一著者名レコードのクラスタリングの結果から、これらの教師付き、教師なしのすべてのモデルの中で、topic-based モデル（著者トピックと論文トピックの関係により著者と論文を結合）が最高の性能であると報告した。Giles のグループの主要な関心は、識別能力の高い論文属性の探求よりも優れたクラスタリングアルゴリズムの発見にあるようであるが、最も著者識別力の高い属性は共著者情報、次いで論文タイトルで、雑誌タイトルはそれより劣るとした。

最近、Cota et al. (2010)が提案した発見的な教師なし方法に基づく階層的クラスタリング法では、融合されたクラスターの集合的情報を、次の段階の融合に利用する。この方法は Giles のグループによる教師付き(Han et al., 2004)、教師なし(Han et al., 2005; Huang et al., 2006)の方法と同じ識別属性（共著者名、論文タイトル、雑誌タイトル）を用いているが、それらと同程度あるいはより優れた性能を示すと彼らは主張している。

McCallum のグループは、著者識別のための書誌レコードの'aggregete'（3つ以上のレコードの組）の比較を可能とし、従って「推移性の破れ」の問題を回避できる表現を検討した(McCallum and Wellner, 2003; Kanani et al., 2007)。この方法を、DBLP からのサンプルレコードのグラフ（ノードは共通の著者名を持つ各レコード、エッジは結合したノード間の属性類似度）の分割に適用した。そして、リンクされたノードのタイトルに含まれる語で Web サーチをしたときヒットすればエッジの重みを増すか、こうして検索された Web ページをノードに加えるかすれば、クラスタリングの性能が著しく向上することを確認した(Kanani et al., 2007; Kanani and McCallum, 2007)。

Torvik のグループは、共通の著者名を含む MEDLINE レコードの対が同一著者によるものである確率を、2つのレコードの間の類似性プロファイルから予測する精緻なアルゴリズムを提案した(Torvik et al., 2005; Torvik and Smalheiser, 2009)。MEDLINE の 2002 年版を用いた大規模な実験(Torvik et al., 2005)により、類似性プロファイルに含まれる属性中最も重要な著者識別情報は共通の共著者の存在であり、次いで発表雑誌の一致、middle name イニシアルの一致であること、この方法で高い再現率(91.9%)と適合率(98.5%)が得られたことを報告した。Torvik and Smalheiser (2009)は更に改良したモデルを MEDLINE の 2006 年版に適用した。種々の観点からモデルを評価した結果、第一著者のフルネームとメールアドレスを識別属性に加えること（これらは MEDLINE レコードだけでなく Web から収集）や、「推移性の破れ」の補正などの改善が高い識別力の達成に効果的であることを明らかにした。

Tang and Walsh (2010)によって提案された ASE (Approximate Structural

---

<sup>10</sup> ここで言う「推移性の破れ」は、著者名 A と著者名 B、及び著者名 A と著者名 C はともに同一著者と判別されたが、著者名 B と著者名 C は同一著者と判定されなかったことを意味する。

## 第2章 先行研究のレビュー

Equivalence)アルゴリズムは、2つの論文に共通の参考文献の情報を論文間の類似度に用い、被引用数の少ない参考文献に高い重みを与えることが特徴である。この情報は同一著者の識別に極めて有効であるが、彼らはこれ以外の識別属性を用いていないので、共通の参考文献のない論文は孤立してしまう。

Smalheiser and Torvik (2009)による総括的レビューでは、ここに挙げた以外のものも含めて、多くの研究で検討された著者識別の方法が報告されている。

## 第3章

# 引用影響要因の可操作化と正確な測定

### 【本章の要約】

論文の被引用数は、論文の質や内容に関わる「内在的」要因だけでなく、質や内容に直接関係はない種々の「外在的」要因にも影響される。3.1 では、これらの内在的・外在的要因を分類するとともに、それらの要因の可操作化、すなわち測定可能な測度(変数)による表現の問題について述べた。

3.2 では、論文の被引用数に影響を及ぼす要因についての先行研究(2.1 でレビュー)の結果を踏まえ、本研究の目的に照らして、被引用数への影響を検討するための外在的要因とそれらを表す測度を選択した。

3.3 では、選択した測度の中で最も測定が困難である論文著者の過去発表論文数について、比較的簡便に正確なデータを得る方法を考案した。本研究で用いる6主題分野の1,395論文に含まれる2,595人の著者に対する著者名サーチから、総計62.9万件の論文が検索された。この中から、同名異人著者による論文(偽論文)を除き、真の著者による論文(真論文)を抽出するため、2段階のフィルタリング処理を行った。一次フィルタリングでは、著者所属アドレスと発表誌の引用関係を利用した。二次フィルタリングでは、検索論文から抽出したサンプル論文に対し目視によって真偽判定を行い、その判定結果を用いて、真偽判定のためのロジスティック回帰モデルを導いた。モデルは分野ごとに設定したが、判別のための説明変数は共通である。このモデルにより、約95%の再現率と適合率、90-95%の正解率を確認した。

## 3.1 論文の被引用数に影響を与える要因とその可操作化

### 3.1.1 内在的要因と外在的要因

論文の被引用数に影響を与える可能性のある要因のうち、その質や内容に係る要因を「内在的」要因(“intrinsic” factors)、質や内容には直接関係のない要因(たとえば著者の数や名声、参考文献数など)を「外在的」要因(“extrinsic” factors)という(Chen, 2012; Peng and Zhu, 2012)。論文の被引用数とその学術的影響度の尺度であるとするならば、それに基本的に影響するのは論文の質や内容に関わる「内在的」要因であろう。しかし、これに対して、質や内容に直接関係はないが、被引用数に影響を与える可能性のある種々の「外在



### 第3章 引用影響要因の可操作化と正確な測定

的]要因も存在する。これらの要因の被引用数への影響に関する先行研究を2.1でレビューしたが、そこでの整理に従い、これらの要因を次のように分類する。

#### (a) 内在的要因

- ①論文の質と重要性：論文の学術的価値、実用(応用的)価値、記述の論理性・明確性など。これらはピアレビューで評価の対象となる項目である。
- ②論文が扱うテーマあるいはトピック
- ③論文が用いている方法：理論か実験か、統計処理方法、症例研究の種別（ランダム化比較試験，コホート研究、症例報告等）など。
- ④結果の有意性の程度
- ⑤研究の学際性・国際性

#### (b) 外在的要因

- ①論文が属する分野
- ②論文のタイプ：原著論文、レビュー、短報など
- ③論文を記述する言語
- ④論文の可視性に関する属性：論文の長さ、図表や数式の使用など
- ⑤著者に関する属性：名声、過去の業績、職位、所属する機関や国、性別、年齢など
- ⑥研究協力に関する属性：共同研究している著者、機関、国の数など
- ⑦参考文献に関する属性：参考文献の数、最新性、重要性など
- ⑧論文が発表された雑誌に関する属性：規模、名声、インパクトなど
- ⑨引用計測期間：被引用数を計測する期間

論文が扱うテーマまたはトピックを内在的要因とし、論文が属する分野を外在的要因としたことについて説明する。ここで「分野」とは、全学問領域を $10^2$ ~ $10^3$ 程度の主題に分類したときの各項目である。引用分析で通常用いられるデータベースである Web of Science や Scopus では、論文単位ではなく雑誌単位に分野（主題カテゴリー）が与えられている。このように大きくくりの分野の中では、論文の内容が多様であるため、これを内在的要因とするには無理がある。なお、後述するように、論文の被引用数は、分野により大きく傾向が異なることが知られている（3.1.2、3.2.1 参照）。一方、論文のテーマやトピックは、分野よりずっと狭い範囲を限定するもので、付与されるキーワード等で示され、論文の特徴を示す属性であるので、ここでは内在的要因とした。

ここでは上記のように内在的要因と外在的要因を分けたが、実は両者を明確に識別することは難しい。外在的要因の中には、内在的要素を含むものもあるからである。たとえば、研究協力に関する属性は、共同研究の拡がりの程度（異なる機関間、国際間）を示すもので、これは内在的であるとも考えられる。また、参考文献に関する属性である参考文献の数やその重要性（参考文献の被引用数等で測られる）は、論文の内容の充実度を示すとも考えられる。従って、これらを単純に外在的要因と言い切ることはできないが、従来の研究において外在的要因に含まれることが多いので、本研究でも外在的要因として扱う。

### 3.1.2 引用影響要因の可操作化

内在的要因にせよ外在的要因にせよ、それらの要因が論文の被引用数に及ぼす影響を定量的に論ずるには、要因という概念を何らかの測度(変数)で表現すること、すなわち可操作化が必要になる。測度は量的な変数の場合もあるし、例えば分野や言語のように質的変数である場合もある。

ある要因をどのような測度で表すかについては、単純に決められる場合もあるが、難しい場合もある。その一つは、要因に対する適切な測度を設定しにくい場合であり、もう一つは、いろいろな測度が考えられて選択に任意性がある場合である。多くの測度が考えられる場合、それらの間に高い相関があれば、どれを採ってもよい訳であるが、そうでない場合は、選択によって異なる結果が得られるかもしれない。このような場合は、独立性の高い複数の測度を説明変数にすることが適当である。

要因の可操作化の難易について、一般的には次のように言える。

#### (1) 内在的要因の可操作化

内在的要因の可操作化は、測度の設定と選択任意性の両面から難しい。最も重要な内在的要因と考えられる論文の質などは、そもそも定量的測度で示すこと自体に無理がある。先行研究においては、この測度はピアレビューでの評点、査読による受理/拒絶、外部研究資金の獲得など、論文から得られるデータとは別の情報源によるものが多い。論文自身から得られる情報に基づいてその質を定量化した試みとして Chen (2012)がある。これは、既存の知識ネットワーク構造を変化させる大きさとして論文の新規性、独創性を定量化した興味ある例であるが、その測度の取得の難度が高いという問題がある。

論文のテーマ、トピックを示す主な測度は、論文に付与されたキーワードである。これは質的変数であるが、キーワードを個別に扱うか、シソーラス等を用いて統合化するか、統計的にクラスター化するかなどの任意性が高い。

被引用数との関係が最近注目されている学際性については、参考文献の分野別多様性を測度とするものが多いが、他にもいろいろな測度が考えられ、やはり任意性が高い。

#### (2) 外在的要因の可操作化

内在的要因に比べると、外在的要因の可操作化は比較的単純である。

最も任意性が高いのは、論文が属する分野の分類であろう。Web of Science を情報源にする場合でも、ESI 分類 (22 分野) と WoS Subject Categories (200 分野以上) のいずれを使うかの選択がある (もちろん別の分類体系を使う選択肢もある)。しかし、どの分類を用いるにせよ、分野が重要な引用影響要因であることは明らかにされている。分野による被引用数格差は、論文あたりの参考文献の数、分野の成長パターン、他分野引用率や他分野被引用率の差の他、データ源として使われる引用索引データベースにおける資料採択率の分野差によってももたらされる (Marx and Bornmann, 2015)。

### 第3章 引用影響要因の可操作化と正確な測定

可操作化が難しいもう一つの外在的要因は、著者の名声あるいは実績である。過去に発表した論文の数や、それらの論文が得た被引用数が測度として用いられることが多いが、論文が共著の場合、(a)第一著者、(b)最も高い業績の著者、(c)全著者の合計または平均、のいずれを選択するかという問題がある。また、過去の論文数の場合は総数か生産密度(たとえば年あたり論文数)か、被引用数の場合は総数か論文あたりの数かの選択がある。更に、論文数や被引用数をカウントするとき、完全計数法(整数計数法)によるか調節計数法(分数計数法)によるかの選択もある。

他の外在的要因を表す測度の選択には、それほど困難さはない。

## 3.2 本研究で用いる要因と測度の選択

### 3.2.1 要因の選択

3.1.1 では、論文の被引用数に影響を与える可能性のある要因を、(a) 内在的要因、(b) 外在的要因に大別し、更にこれらを細分した。1.2 の冒頭で述べたように、本研究の目的は、外在的要因が論文の被引用数にどの程度影響を与えるかを明らかにすることなので、(a)の内在的要因は対象としない。(b)の外在的要因については、次のようにする。

- (1) 論文が属する分野 (①) によって被引用数の傾向が大きく異なることを 3.1.2 の(2)で述べた。従って、異なる分野が混在した論文集合を用いて引用分析を行う場合は、分野をダミー変数とする、あるいは分野差を考慮して被引用数を規格化する(Waltman and van Eck, 2013a, 2013b)などの方法が採られる。しかし、より明確に分野間の比較ができるようにするため、本研究では、複数の分野別に分析を行い、分野共通の傾向を見出す方法を採用。具体的に選定した分野は、Web of Science の 200 以上の主題カテゴリーから選んだ 6 つのカテゴリーで、1.4.1.1 に述べたとおりである。
- (2) 論文のタイプ (②) と論文を記述する言語 (③) によっても被引用数は大きく異なる。論文タイプの場合、正規の原著論文と比べレビュー論文はよく引用され、短報は引用が短期的である。また、当然のことながら、英語の論文は多言語の論文に比べ引用される可能性が高い。本研究では、雑誌論文として代表的である英語の原著論文を対象とする。
- (3) 外在的要因の④～⑧をすべて対象とし、これらの要因による被引用数への影響について第4章で検討する。
- (4) 引用計測期間 (⑨) の被引用数への影響については、他の要因とは異なるアプローチが必要であるので、別途に第5章で検討する。

### 3.2.2 各要因を表す測度の選択

### 第3章 引用影響要因の可操作化と正確な測定

3.1.2 で見たように、ある要因を表す測度は一様ではない。本研究では、原則として1つの要因から少なくとも2つの測度を選択することとした。これは、同じ要因に属する測度でも、被引用数に与える影響のパターンは異なることがあるためである。但し、雑誌に関する属性(⑧)については、後述の理由から1測度とする。

具体的には、各要因から以下の測度を選択する。

#### 3.2.2.1 研究協力に関する要因

ほとんどの先行研究で取り上げられている著者数に加えて、所属機関数、所属国数の3つを測度として選択する。単独著者か複数著者か、機関内共著/同一国内異機関間共著/国際共著かという質的測度を用いた研究もあるが、前述の3測度はこれらの区別を定量的に表すことができる。

#### 3.2.2.2 参考文献に関する要因

多くの先行研究で取り上げられている参考文献数と、適用例は多くないがその最新性を、測度として選択する。後者については、参考文献のうち、その発表年が論文の発表年の5年前以内のもののパーセンテージ(Price指数)を具体的な測度とする。

参考文献のインパクト(参考文献の平均被引用数等で測られる)は、少数の研究で論文の被引用数との強い関係が示唆されているが、論文の直接の属性とは言いがたいこと、発表された雑誌のインパクトとの相関が高いと予想される(インパクトの高い雑誌の論文は、やはりインパクトの高い雑誌の論文を優先的に引用する)ことから、選択しなかった。

#### 3.2.2.3 論文の可視性に関する要因

多くの研究で取り上げられている論文の長さの他に、論文中の図の数、表の数、数式の数に加え、4つの測度を選択した。図の数、表の数は少数の先行研究で取り上げられているが、被引用数との有意な関係は示されていない。また、数式の数も測度とした先行研究は恐らく存在しない。しかし、これらはいずれも科学論文の重要な構成要素であり、他の多くの測度とともにその影響を検討する価値があると考えた。

いくつかの研究で、タイトルや抄録の特性(長さ、可読性、文型など)と被引用数の関係が議論されているが、具体的測度の選び方、測定の方法に任意性が高いため、用いないことにした。

#### 3.2.2.4 著者に関する要因

本研究で用いる測度は、すべて著者の過去の業績に関するものであり、次の3つである。

- (a) 生産性：対象の論文を発表するまでの発表論文数で表す。
- (b) 引用影響度：(a)で示した過去論文が対象論文発表までに得た被引用数で表す。
- (c) 活動期間：著者の最初の論文と対象論文の発表年の差で表す。

### 第3章 引用影響要因の可操作化と正確な測定

これらのデータは対象の論文から得ることはできず、著者の過去の論文を検索・特定し、更にそれらの論文が得た引用数を調べる必要がある。従って、これらの測度の測定は他の測度に比べて著しく困難であり、かつ誤差も生じ易い。それにも拘わらずこれらを選択したのは、いくつかの先行研究（特に2.3で述べた総合的研究）において、これらの測度の重要性が指摘されていることによる。

著者の所属機関に関する特性（所属機関が大学かそれ以外か、名声の高い機関か否かなど）や所属国に関する特性（米国等特定の国や地域に属しているかなど）は、過去実績に比べれば測定が容易であるが、定量的測度ではないこと、3.2.2.1の所属機関数や所属国数と組み合わせる意味が低いと考えられることから、選択しないこととした。

選択した3測度(a)、(b)、(c)をどのような量として測定するかについては、3.1.2(2)で述べたように多くの選択肢がある。本研究では、論文の第一著者の過去論文数、それらの論文が得た被引用数、活動期間を用いることとし、論文数と被引用数は調節計数法によって求めた。また、論文数と被引用数には、累積実績測度と効率実績測度の両方の適用を試みた。詳細については第4章で述べる。

#### 3.2.2.5 雑誌に関する要因

この要因を表す測度で代表的なものは雑誌の引用インパクト指標であり、その中でもインパクトファクター(JIF)がよく用いられている。そして、これらを用いたほとんどの研究で、論文への最も強い影響要因の一つであると報告されている。

しかし、第4章の研究では、雑誌の引用インパクトを測度として用いず、代わりに、論文が発表された雑誌をダミー変数として含めることにした。その理由は、標本データとする論文を、それぞれの分野から選択した4種の雑誌の論文に限定したことによる。このような選択をすれば、雑誌に関する測度は4つの値のいずれかに限定されることになるので、特定の測度を用いるよりも、雑誌をダミー変数とすることにより、いろいろな雑誌特性がダミー変数値に吸収されると考えた。なお、第5章の研究では、引用影響度とJIFの間に興味ある関係が想定されたため、JIFを測度とした。

各要因に対して選択した測度を表3-1にまとめた。

表 3-1 本研究で用いる引用影響要因の測度

要因	測度	注記
分野	なし	6つのWeb of Science主題カテゴリー(分野ごとに分析)
論文タイプ	なし	原著論文に限定
言語	なし	英語の論文に限定
研究協力要因	著者数 所属機関数 所属国数	
参考文献要因	参考文献数 Price指数	
論文可視性要因	論文の長さ 図の数 表の数 数式の数	
著者要因	第一著者の過去発表論文数 第一著者の上記論文が得た被引用数 第一著者の活動年数	
雑誌要因	雑誌(名義尺度)	

### 3.3 測度の正確な測定－著者の過去実績データの例

論文の被引用数への影響を検討する外在的要因に対して 3.2 で選択した測度の多くは、論文自体から測定することができ、測定法にそれほど大きな問題はない。もちろん、測度の定義にそれぞれある程度の任意性があり、測定誤差を生ずる可能性もあるが、多くの測度に関しては、巨視的な統計的研究に重大な影響を及ぼす可能性は低い。

しかし、3.2.2.4 で述べたように、著者の過去実績に関する測度のデータは対象の論文から得ることはできず、測定が著しく困難でかつ誤差も生じ易い。そこで本節では、これらの測度データを、比較的簡便な方法で正確に測定するために採った方法について述べる。

なお、本節(3.3)に述べる内容は以下の論文が核となっているが、この論文の共著者 1 名から、インターネット公表に対する同意を得ることができなかつた。このため、一部を除くあるいは短縮して公表する。詳しい内容については下記論文を参照されたい。

Onodera, N. et al. A method for eliminating articles by homonymous authors from the large number of articles retrieved by author search. *Journal of the American Society for Information Science and Technology*, 2011, vol. 62, no. 4, p. 677–690.

### 3.3.1 著者論文数データ測定の概要

本研究で用いる著者実績に関する測度は、過去の論文数、それら論文が得た被引用数、及び活動期間であるが、測定の基本となるのは過去の論文数である。過去に発表した論文が明らかになれば、それらが得た被引用数はWeb of Scienceから得ることができる。また、活動期間は、基になった論文と最も早期の論文の発表年の差とすればよい。そこで以下では、過去論文数の測定について述べる。

このデータを得るには、通常、データベースを用いて著者名による論文検索を行う。しかし、その際に、同姓同名の異なる著者の論文の混入が大きな問題になる (Moed, 2005, Chapter 14; Aksnes, 2008)。

同名著者の論文集合を、それぞれの著者の論文に分割する方法 (著者識別方法という) は種々開発されており、2.3 にその主なものをレビューした。その多くは、異なる同名著者による書誌レコードを個々の著者ごとのクラスターに分割する方法を採るが、そのためには、原則として集合中のすべてのレコード対間の類似度を計算しなければならない。一方ここでの目標は、著者名サーチで得られたある検索論文が、意図している著者によるもの (以下「真論文」という) か同名の異なる著者によるもの (以下「偽論文」という) かを識別することであるので、すべての検索論文間の類似度を比較する必要はない。このため我々は、検索論文集合をクラスター分割するのではなく、より簡便に、個々の検索論文をもとの論文と比較するアプローチを採ることとした。この意味で、ここでの狙いは Wooding et al. (2006)のそれに似ているが、彼らよりずっと大きい論文集合を扱っており、主題分野、所属機関の広がりも大きい。

目標としては、同定洩れ (真論文を偽論文と判別) を多少犠牲にしてもノイズ (偽論文を真論文と判別) を最小限に抑えることとした。クラスター分割の方法を採らないため「推移性の破れ」の問題 (2.3 の脚注 10 を参照) を考慮に入れていないという限界を考慮し、少なくとも 90%の正解率を得ることを目標とした。

### 3.3.2 著者識別対象の論文集合

第4章及び第5章で、被引用数及び引用持続性を説明するための重回帰分析にサンプルとして用いるのは、6つの分野に属する24の雑誌 (分野ごとに4誌) から抽出した2000年発表の正規論文1,395件 (1誌から60以内) である。これらの分野、雑誌、論文の選定については1.4.1に詳述した。

これら24誌のうち6誌 (各分野から1誌) の論文からは全著者を、他の18誌の論文からは第一著者を選んで、論文の著者が過去に発表した論文の数を調査した。調査した著者は延べ2,595名となった。

以下この節(3.3)では、これらの24の雑誌を「ソース誌」、1,395の論文を「ソース論文」、

## 第3章 引用影響要因の可操作化と正確な測定

2,595名の著者を「ソース著者」と呼ぶ。

それぞれのソース著者について、WoSを用いて著者名サーチを行った。検索は、ソース論文の発表年である2000年までの期間について行った。使用したWoSは1970年まで遡及できるので、ソース著者のほとんどの活動期間をカバーしていると考えられる。

検索対象の期間におけるWoSの著者名データは、原則として姓+first name イニシアル(たとえば“Smith, A”)または姓+first name イニシアル+second name イニシアル(たとえば“Brown, AB”)の形で表現されている<sup>11</sup>。検索の著者名は、WoSでもとのソース論文において表記されている通りの綴りを用いた。

2,595のソース著者名に対して、総数62.9万の論文が検索された(ソース論文自身を除く)。以下ではこれらを「検索論文」と呼ぶ。これは1ソース著者あたり242論文に当たる。この数字は、一人の研究者の最長30年間(1970–2000)における平均的な生産論文数を遙かに超えており、明らかに大量の偽論文が混入している。

### 3.3.3 真論文と偽論文を判別する方法

この章では、これらの検索論文を真論文と偽論文に判別するための手続きを示す。

#### 3.3.3.1 判別に用いる論文属性情報

真論文と偽論文の識別のために利用する情報は次の通りである。括弧内は、それぞれの属性に対する変数名である。

- (1) ソース論文と検索論文に共通する共著者の存在(*AutMatch*)
- (2) ソース論文と検索論文の著者所属機関アドレスの類似度(*Add\_Sim*)
- (3) ソース論文誌と検索論文誌の引用関係の強さ(*X*)
- (4) ソース論文と検索論文のタイトル語の類似度(*Tit\_Sim*)
- (5) 検索論文がソース論文に引用されていること(*Cited*)
- (6) ソース論文と検索論文の間の共引用の存在(*Cocit*)
- (7) ソース論文と検索論文の発表年の間隔(*Age*)
- (8) ソース著者の所属国が特定の国(中国、台湾、韓国、日本)であること(*FEA*)

#### 3.3.3.2 一次フィルタリングのためのデータ前処理

ここでは、前節で挙げた8つの判別属性のうち5つについて、変数を数値化する方法を示す。62.9万の検索論文のすべてに、これらの属性値を付与する。他の3属性の数値化については3.3.3.4で述べる。

---

<sup>11</sup> 2006年以降の論文の著者名データには、多くの場合 full first name が付与されるようになった。



### 第3章 引用影響要因の可操作化と正確な測定

#### (1) ソース論文と検索論文に共通する共著者の存在

WoS の著者名記述で少なくともひとつが両論文間で合致した場合、その検索論文の *AutMatch* の値を 1、そうでない検索論文ではその値を 0 とする。

#### (2) ソース論文と検索論文の所属機関アドレスの類似性

全検索論文とソース論文からすべての著者所属機関アドレスデータを取り出し、その中に含まれる各単語（国名部分の単語は除く）に、出現頻度によりウェイトを与える。すべての検索論文について、そのソース論文と所属機関アドレスを照合し、両方のアドレスの国名が同じ場合に、国名を除いた部分に共通に含まれる単語のウェイトの和を *Add\_Sim* の値とする（国が一致するアドレスがないときの *Add\_Sim* 値は 0 である）。

#### (3) 雑誌の引用関係の強さ

Journal Citation Reports (JCR) Science Edition の 2004 年版を用いて、2000–2004 の 5 年間に、ソース論文雑誌が検索論文雑誌を引用した回数と検索論文雑誌がソース論文雑誌を引用した回数を数え、その平均を  $X$  とする。

#### (4) ソース論文と検索論文の発表年の間隔

検索論文の発表年を  $y$  とすればその *Age* は  $2000-y$  である。

#### (5) ソース著者の所属国が特定の国であること

ソース著者の所属が日本、中国、韓国、台湾のいずれかであるとき、その検索論文の *FEA* の値を 1、そうでないとき 0 とする。

#### 3.3.3.3 一次フィルタリング

第一段階では以下の手順で偽論文の除去を行った。

- 1) *AutMatch* = 1 の検索論文は真論文である可能性が高いので、以下の除去基準に拘わらず二次フィルタリングに送る。
- 2) *Add\_Sim* が 5 未満の検索論文を除去する。
- 3)  $X$  の値が 0 の検索論文を除去する。

#### 3.3.3.4 二次フィルタリングのためのデータ前処理

3.3.3.2 で説明しなかった 3 つの属性の数値化方法を述べる。これらの属性値は、上述の一次フィルタリング段階を通過した検索論文に対してのみ計算する。

#### (1) ソース論文と検索論文のタイトル語の類似性

6 つの分野のそれぞれに対して、予備的ロジスティック回帰によって「真」と予測された検索論文から 2,000 を無作為抽出する。これとすべてのソース論文（分野ごとに約 240 論文）を合わせたコーパスを分野ごとに作り、その論文タイトルから抽出した単語を整理（ストップワード除去、語尾処理等）した後、逆文書頻度 (*idf*) に基づいてそれぞれの分野における各語の重みを定める。その後、すべての検索論文について、タイトルから抽出された語をそのソース論文から抽出された語と比較し、マ

### 第3章 引用影響要因の可操作化と正確な測定

ッチした語の重みの和を、その検索論文の *Tit\_Sim* の値とする。

#### (2) ソース論文による検索論文の引用

検索論文がソース論文に引用されているときその *Cited* の値を 1、そうでない場合 0 とする。

#### (3) 検索論文とソース論文の共引用の存在

検索論文を引用した論文とそのソース論文を引用した論文（いずれも 2006 年までの引用）の間に、共通な論文が少なくとも 1 つ存在するとき、その検索論文の *Cocit* の値を 1、そうでない場合 0 とする。

#### 3.3.3.5 二次フィルタリング

一次フィルタリングを通過した検索論文から更に偽論文を除くため、次の手順により二次フィルタリングを行った。

##### (1) サンプル論文の目視による判別

6つの分野からそれぞれ 500 の検索論文(計 3,000 論文)をサンプル論文として抽出し、それぞれのサンプル論文の真偽を 2 名の判定者が目視判定した。2 名の判定者の判定が一致しなかった論文(全サンプル論文の 11.2%)は別の 1 名の判定者の目視により決定した。

##### (2) ロジスティック回帰による判別関数モデル化

サンプル論文の目視判定の結果に基づき、検索論文が偽論文である確率  $p$  を予測するロジスティック重回帰モデルを導出した。回帰モデルは 6 つの分野別に設定した。

ロジスティック回帰の従属変数 *Judge* の観測値は、目視判定が「真」なら 0、「偽」なら 1 である。3.3.3.2 と 3.3.3.4 で述べた 8 つの独立変数を考えたが、*Cited* と *Cocit* のどちらかの値が 1 であるサンプル論文ではすべての場合に目視判定が真であるため、真偽判定を次の 2 ステップで行うこととした。

1) *Cited* = 1 または *Cocit* = 1 であるサンプル論文は、他の属性値に拘わらず真論文と判別する。

2) 残りのサンプル論文(*Cited*=0 かつ *Cocit*=0)に対して 6 つの独立変数を用いたロジスティック回帰を行う

##### (3) 全検索論文の真偽判別

まず、*Cited* = 1 または *Cocit* = 1 の検索論文はすべて「真」と判別した。次に、(2) で得られたロジスティック回帰モデルを用いて、一次フィルタリングを通過した残りの全論文の真偽を判別した。判定の境界を  $p=0.5$  とした。

### 3.3.4 結果と考察

#### 3.3.4.1 一次フィルタリング

62.9 万件の検索論文のうち 106,163 論文が一次フィルタリングを通過した (1 ソース著

者あたり 40.9 論文)。これは当初の全検索論文の 16.9%である。106,163 論文中 39,239 論文 (37%)は *AutMatch* = 1 の論文である。

#### 3.3.4.2 二次フィルタリングのためのサンプル論文の判別

##### (1) サンプル論文の目視による判別

人間の判定者による目視判定の結果、全てのサンプル論文のうち 75%が真論文と判定された。*AutMatch* = 1 のサンプル論文のほとんどは真と判定された。これに対して、*AutMatch* = 0 のサンプル論文では、真と判定された論文の比率は分野により大きく異なる。これは主に識別属性 (特に *FEA*) の分布の分野間の違いによるもので、これらの属性値がほぼ同じ条件で比較すると、各分野の真論文率に大きな差は見られなかった。

##### (2) 予測に使われた個々の独立変数と判定結果の関係

独立変数の真論文判別のための予測力を予備的に検討するため、個々の独立変数と目視判定の結果の関係を調べた。

*Add\_Sim*、 $\log(X+1)$ 、*Tit\_Sim* の 3 つの変数は、それらの値の増加に連れて真論文率が上昇する。日中韓台の所属著者の論文 (*FEA* = 1) の真論文率は、*FEA* = 0 の論文に比べて非常に低く、*FEA* がよい予測因子であることを示唆する。論文の年齢 (*Age*) は、単独では真論文率との相関はほとんどなかった。3.3.3.5 の(2)で述べたように、*Cited* = 1 または *Cocit* = 1 であるサンプル論文はすべて真論文と判定された。

##### (3) 独立変数間の相関

6 つの独立変数間の相関係数を各分野で求めた。最も相関が強い場合でも相関係数の絶対値は 0.5 程度で、変数間に著しく強い相関は見られず、多重共線性の問題は低いと思われる。このため、以下のロジスティック回帰分析には 6 つの独立変数すべてを含めた。

##### (4) 真偽判別のためのロジスティック重回帰モデルの検証

各分野において、3.3.3.5 で述べたロジスティック回帰モデルに従って真偽判別を行った。*Cited* = 1 または *Cocit* = 1 であるものを除いたサンプル論文を無作為に訓練群 (各分野約 300) と検証群 (各分野約 130) に分け、訓練群に対して分野ごとに回帰分析を実行した。

*AutMatch*、*Add\_Sim*、*Tit\_Sim* の 3 変数は全ての分野で有意な予測変数であり、 $\log(X+1)$ 、*FEA* も若干の分野を除けば有効である。*Age* の予測力は他の予測変数に比べてやや弱いですが、それでも 2 分野で 5%水準有意である。それ故、変数の選択は行わず、真論文予測の重回帰モデルにはこの 6 つの独立変数を含めた。

訓練群に対して得られた回帰モデルを用いて、訓練群と検証群両方のサンプル論文の真偽判別を行った。回帰モデルから予測される *p* 値が 0.5 未満のとき真論文、0.5 以上のとき偽論文とした。この判別を人間の目視による判定と比較した結果、真論文の 95%は正しく真論文と判定されたが、誤論文を正しく誤論文としたのは 80%強にとどまった。適合率は 95%前後、正解率は 90-95%であった。全体的に見て、我々が当初目標とした正解率 90%以上を達成することができた。

## 第3章 引用影響要因の可操作化と正確な測定

### (5) 全検索論文の判別のための回帰モデルの決定

訓練群で得られた回帰モデルが訓練群のみならず検証群にもよく適用されることが確認されたので、全サンプル論文 (*Cited*=1 または *Cocit*=1 のものを除く) を用いて、最終的な回帰モデルを分野ごとに決定した。これによる性能は訓練群に対して求めたものとほぼ同様で、再現率、適合率、正解率はそれぞれ 95%、93%、92%であった。

#### 3.3.4.3 二次フィルタリングにおける真論文判別

一次フィルタリングを通過した 106,163 検索論文のうち、12,947 論文(12.2%)は、ソース論文に引用されているかソース論文と共引用されている (*Cited*=1 または *Cocit*=1) ため、真論文と判定された。それ以外の 93,216 の検索論文について、3.3.4.2 の(5)で求めた各分野のロジスティック回帰式により真偽判別を行った。

真論文と判定されたのは 90,052 論文で、一次フィルタリングを通過した 106,163 論文の 84.8%に当たる。*Cited*=1 または *Cocit*=1 のため真論文と判定された論文を除くと、93,216 論文中 77,105 論文がロジスティック回帰によって真と判定されたので、通過率は 82.7%である。

### 3.3.5 まとめ

#### 3.3.5.1 得られた結果の要約

2,595 人のソース著者に対して、当初の著者名サーチで得られた論文は約 62.9 万件であったが、そのうち一次フィルタリングを通過したのは 10.6 万件(16.9%)、二次フィルタリングを通過して最終的に真論文と判定されたものは 9.0 万件(14.3%)であった。一次フィルタリングで大量の論文が除かれたが、そこで残った論文の大部分が二次フィルタリングを通過した。一次フィルタリングでは偽論文を除去して論文を絞り込むことに重点を置いたため、除去された論文に相当の真論文を含む可能性がある。すなわち、ソース論文と全くアドレスが異なっているため、あるいは雑誌の引用関係が全くないため一次フィルタリングで除かれた論文にも、真論文がある程度存在すると考えられる。

サンプル論文の検証結果から、二次フィルタリングに用いたロジスティック回帰モデルによって 90%以上の正解率が得られることが実証された。但し、この性能は一次フィルタリングを通過した論文のみに対して得られたものである。

同名の共著者の存在及びアドレスの類似度は最も重要な判別要素であり、タイトル語の類似性と雑誌間の引用関係強度も判別に有効であった。また、特定の国(日・中・韓・台)の著者に同名異人が極めて多いため、著者がそれらの国に所属するか否かを判別式に組み込むことが重要であった。ソース論文に引用されているかソース論文と共引用されている検索論文は、ほぼ確実に真論文であることが判った。それ故、これらの情報は判別に極めて有効である。

#### 3.3.5.2 先行研究との比較

ここで提案した方法では、WoS データベースから入手可能な属性をできるだけ多く利用した。それらは、*AutMatch*、*Add\_Sim*、*Tit\_Sim*、*Age*、*FEA* のような書誌データに基づく属性だけでなく、*log(X+1)*、*Cited*、*Cocit* のような引用データに基づく属性も含む。一方、これまでに報告されている多くのアプローチは、Giles のグループ(Han et al., 2004, 2005; Huang et al., 2006; Song et al., 2007)、McCallum のグループ(McCallum and Wellner, 2003; Kanani et al., 2007; Kanani and McCallum, 2007)、Cota et al. (2010) など、より優れたクラスタリングアルゴリズムを開発することに主眼を置いているため、判別に用いた属性は比較的少ない。

これに対して、Torvik のグループ(Torvik et al., 2005; Torvik and Smalheiser, 2009)はこの両方の目的を追求している。彼らの結果は我々の結果とある程度一致している。但し、我々は second name イニシアル、MeSH 索引語、メールアドレス、full first name の情報は用いておらず、雑誌名の代わりに雑誌間の引用関係を用いた。一方、我々は Torvik らの用いていない引用、共引用の情報を用い、これらが極めて価値の高い著者識別情報であることを示した。Torvik のグループが提案したアプローチは、厳密なクラスタリング法を用いるとともに多くの判別属性を採り入れることにより、非常に高い著者識別能力を示す。これに対してここで採ったアプローチ（個々の検索論文をそのソース論文と比較）は、「推移性の破れ」の問題を解決できないなどいくつかの限界があるが、大量のデータを比較的少ないステップで処理できるという利点がある。それ故、提案した方法は、我々が直面した状況—大量の同名異人著者による「偽」論文が混入した論文集合から「真」論文をある程度（完全ではないが）の精度で識別する—には適していると考えられる。

#### 3.3.5.3 提案した方法の拡張可能性

6 つの異なる分野にわたり、多様な所属を持つソース著者に対して共通の手法（同じ独立変数によるロジスティック回帰分析）を適用し、得られた結果（偏回帰係数の値と有意レベル、及び判別の性能）は分野間でそれほど大きくは違わなかった。このことは、多くの分野と多様な所属の著者に広くこの方法が適用可能であることを示唆する。ここでは WoS からのデータにこの方法を適用したが、共著者、著者の所属アドレスと所属国、タイトル語、発表年に関するデータは他の書誌データベースでも同じように得ることができるので、これらの属性に関する結果は一般化できる。一方、判別属性とした *log(X+1)*、*Cited*、*Cocit* の計算に用いられる引用データは、WoS など引用索引を持つ少数のデータベースからしか得ることができない。引用に関する属性が著者識別に有効であることを実証的に示した研究は、書誌結合情報を用いた Tang and Walsh (2010) の研究以外ほとんど見られないので、この種の属性の有効性を示したことは本研究の特徴のひとつと言える。

## 第4章

# 論文の被引用数に対する外在的要因の影響

### 【本章の要約】

論文の被引用数に影響を与える要因に関して、異なる主題分野にわたって何らかの共通の傾向が見られるか否かを検討した。特に、論文の質や内容に直接関係しない「外在的」要因の影響に注目した。6つの分野（物性物理学、無機・核化学、電気・電子工学、生化学・分子生物学、生理学、消化器病学）のそれぞれから、同じ年（2000年）に発表された230-240の原著論文を抽出した。それらの論文が発表後6-7年間及び11-12年間に得た引用をカウントし、その被引用数を負の2項重回帰(NBMR)モデルにより予測した。NBMRの説明変数には、著者の研究協力、参考文献、論文の可視性、著者の過去実績（発表論文数、それらの論文の被引用数等）、発表した雑誌等、種々の属性を含めた。選択された予測変数、及びそれらの有意性の強さには、分野を越えた一般的傾向が見られた。最も強い予測変数はPrice指数、次いで参考文献数であった。著者数や著者実績の効果は比較的弱かった。

### 4.1 本章における研究目的

1.1において、研究評価に計量書誌学データ、特に引用データを用いることについての意味、限界、注意を要する点について述べた。その要留意点のひとつは、論文の被引用数とその質や内容に直接関係のない様々な要因(外在的要因)の影響を受けることであるが、これについての知見を深めるため、1.2において3つの研究目標を設定した。本章は、そのうちの「研究目標2」について論述するものである。

研究目標2：論文の被引用数に影響を及ぼす種々の外在的要因の影響を体系的に検討し、どの要因がどの程度の影響力を持つか分析する。複数の分野についてこれを行うことにより、影響を与える主要因について分野共通の傾向があるかどうかを明らかにする。

被引用数に影響を及ぼす要因について、多くの研究で検討されていることを2.1で述べた。特に、2.1.3で述べた研究では、重回帰分析等により種々の要因の影響を総合的に検討しており、それぞれ興味深い結果を得ている。しかしこれらのいずれも、標本とした論文が、比較的狭い分野やトピック(Stewart, 1983; Peters and van Raan, 1994; van Dalen

#### 第4章 論文の被引用数に対する外在的要因の影響

and Henkens, 2001, 2005; Walters, 2006; Haslam et al., 2008; Lokker et al., 2008; Fu and Aliferis, 2010; Chen, 2012; Peng and Zhu, 2012; Didegah and Thelwall, 2013)、特定の雑誌(Bornmann and Daniel, 2008; Davis et al., 2008)、あるいは特定の国の論文(He, 2009)に限定されているため、それらの結論に汎用性があるとは言えない。また、Bornmann and Daniel (2008)、Davis et al. (2008)、He (2009)、Chen (2012)は、種々の要因を総合的に検討するというより、ある要因と被引用数の関係に焦点を置いたもので、他の要因は制御変数として追加されている程度である。

これに対して、本研究は、複数の分野に対して同一の外在的要因に関する測度を用いて、被引用数（及び引用持続性）に影響を与える要因について体系的に分析を行い、分野を越えた一般的傾向が見られるかどうかを考察するものである。

第3章で、論文の引用に影響を及ぼす可能性のある外在的要因と、それらの要因を表す測度について考察した。そして、本研究で対象とする要因と測度を3.2において選択し、表3-1にまとめた。そこで述べたように、被引用数の傾向が分野、論文タイプ、言語に依存することはよく知られているので、本研究ではそれ以外の外在的要因に着目する。このため、いくつかの異なる分野について、同じ年にいくつかの雑誌に発表された同じ言語・タイプの論文（英語の原著論文）の間の被引用数の違いと、種々の要因へのその依存性を検討する。同じ年に発表された同じ分野、同じ論文タイプ、同じ言語の論文の中で被引用数を比較することにより、これらの影響は除去され、それ以外の外在的要因の影響、すなわちどんな要因がどの程度の影響を与えるのかについて分析することができる。また、複数の分野についてこの分析を行うことにより、影響を与える主要因について分野共通の傾向があるかどうかを明らかにすることができる。

被引用数への種々の要因の正味の寄与（他の要因の影響を除去した寄与）を評価するために負の2項重回帰分析を用いる。これらの要因には論文の質や内容に関する特徴は含まないので、得られるモデルは被引用数をある程度予測できたとしても、高精度の予測は期待できない。この研究の目標は説明力の高いモデルを得ることではなく、外在的な要因から期待される被引用数のベースラインを得ることにある。分野を越えてある程度共通性の高いベースラインを得ることができれば、実際の論文のこのベースラインからの乖離は、その論文のインパクトをこれまでより適切に測る指標となり、研究評価に引用データがより正確に応用できるかもしれない。

以上の考えに立って、ここでは次の戦略を採ることとする。

- (1) 複数の分野からそれぞれいくつかの雑誌（すべて英語のみ）を選ぶ。
- (2) 各雑誌から同一年に発表された原著論文を抽出する。
- (3) 分野ごとにいくつかの負の2項重回帰モデルを比較検討する。
- (4) その目的変数は、論文が得た被引用数とする。
- (5) その説明変数には、被引用数に影響を与えると想定される多くの測度（表3-1に示したもの）を含める。

## 第4章 論文の被引用数に対する外在的要因の影響

(6) 雑誌の引用インパクトの影響を考慮して、雑誌をダミー説明変数とする。

これによって、被引用数に対する潜在要因の影響をそれぞれ分離し、各要因の寄与を評価することができる。

## 4.2 データ源と方法

### 4.2.1 対象とする分野及びサンプル論文

ターゲットとする分野として、Web of Science（以下 WoS と略記）の主題カテゴリー (Subject Categories) から次の 6 つを選んだ。以降では、これらの分野を下記の下線部で略記する。

- ・ 物性物理学(condensed-matter physics)
- ・ 無機・核化学(inorganic chemistry)
- ・ 電気・電子工学(electric and electronic engineering)
- ・ 生化学・分子生物学(biochemistry and molecular biology)
- ・ 生理学(physiology)
- ・ 消化器病学(gastroenterology)

そして、これらのそれぞれの分野から、分析用の論文を抽出する雑誌を 4 誌ずつ選択し、各雑誌の 2000 年発表の原著論文 (WoS の記事タイプが "articles" であるもの) から、それぞれ 50-60 件を無作為抽出して、重回帰分析のためのサンプル論文とした。

このように分野、雑誌、サンプル論文を選択した理由については、1.4.1 に詳しく述べた。選択した 24 の雑誌 (6 分野から 4 誌ずつ) と、それらにおける 2000 年に発表された正規論文数及び抽出したサンプル論文の数は表 1-1 に示した通りである。

### 4.2.2 被引用数データの取得

これらのサンプル論文が得た被引用数のデータを、2006 年 10 月と 2011 年 12 月に WoS から取得した。従って、引用ウィンドウの長さはそれぞれ 6-7 年、11-12 年になる。以下では、これらの 2 つの引用ウィンドウに相当する被引用数を負の 2 項重回帰分析の目的変数とし、それぞれ  $C_6$ 、 $C_{11}$  と表記する。

これらの被引用数には自己引用を含むが、1.4.2.1 に述べたように、本章での研究のように比較的長い引用ウィンドウを用いる巨視的統計解析では、自己引用が大きなバイアスをもたらす可能性は小さいと考えられる。

### 4.2.3 被引用数に対する説明変数のデータの取得



## 第4章 論文の被引用数に対する外在的要因の影響

3.2 で選択し、表 3-1 に示した測度を、負の 2 項重回帰分析の説明変数とする。以下に、これらの測度に対する説明変数としての名称を示す。

- ・研究協力に関する要因
    - (a) *Authors* : 論文の著者数
    - (b) *Institutions (Insts)* : 著者の所属する機関(著者所属機関)の数
    - (c) *Countries* : 著者所属機関が属する国(著者所属国)の数
  - ・参考文献に関する要因
    - (d) *References (Ref)* : 論文に引用された参考文献の数
    - (e) *Price* : Price 指数 (参考文献のうち、その発表年が論文の発表年の 5 年前以内のものパーセンテージ).
  - ・論文の可視性を示す要因
    - (f) *Figures* : 論文中の図の数
    - (g) *Tables* : 論文中の表の数
    - (h) *Equations (Eqs)* : 論文中の番号付き数式の数
    - (i) *Length* : ページあたり文字数を規格化した論文のページ数
  - ・著者の過去の業績を示す要因
    - (j) *Published articles (Publ)* : 論文の第一著者により 2000 年までに発表された論文数
    - (k) *Cited* : ‘*Publ*の論文が 2000 年までに得た引用の総数
    - (l) *Age* : 第一著者の活動期間 (‘*Publ*の論文中最初の論文の発表年から 2000 年までの経過年)
    - (m) *Rate of publication (RatePubl)* : 第一著者の活動機関内の年平均発表論文数(=*PublAge*)
    - (n) *Median of the number of citations (MedCites)* : ‘*Publ*の各論文の年あたり被引用数 (2000 年まで) の中央値
- 以下では、このうち(j), (k), (l)をまとめて累積実績指標、(m), (n)をまとめて効率実績指標という語を用いる。
- ・論文が発表された雑誌の要因
    - (o) *Jnl1; Jnl2; Jnl3* : 論文が発表された各雑誌を表すダミー変数

個々のサンプル論文に対するこれらの属性の値を、次の手続きによって得た。

*Authors*, *Insts*, *Countries*, *Refs*, *Price* のデータは、ダウンロードした WoS レコードから得られる。*Authors* は AU フィールド、*Insts* は C1 フィールドのエントリー数を数えた。*Countries* は C1 フィールドの各エントリーの末尾の国名記載より得た<sup>12</sup>。*Price* は、

<sup>12</sup> 米国のアドレスは末尾が州名である場合があり、これらをすべて USA に統合した。England, Scotland, Wales, Northern Ireland は UK に統合し、Hong Kong は China に統合した。

## 第4章 論文の被引用数に対する外在的要因の影響

CR フィールドの参考文献リストから、発表年が 1996-2000 年（論文発表から 5 年以内）のものを数えることにより求めた。

*Figures, Tables, Eqs* は、オリジナルの論文から直接カウントした。*Figures* には、チャート、ダイアグラム、写真等を含む。*Length* については、それぞれの雑誌からのサンプリングにより 1 ページあたりの平均文字数を求め、標準的文字数（6,400 文字/ページ）に換算したページ数を各論文の値とした。

著者の過去実績に関する属性(*Publ, Cited, Age, RatePubl, MedCites*)のデータの取得は、同名異人著者の存在のため難しい問題であった。1,395 件のサンプル論文のすべての著者（延べ 6,000 人近く）についてこれらのデータを得るのは時間的に困難であったため、全サンプル論文の第一著者（すなわち 1,395 人）のデータを用いることとした<sup>13</sup>。まず、3.3 に述べた方法により、*Publ* の値（各第一著者が 2000 年までに発表した論文数）を得た。その中の最も古い論文の発表年と 2000 年の差を *Age*（第一著者の活動期間）とし、*Publ* を *Age* で割った値を *RatePubl* とする。*Cited* と *MedCites* の算出に必要な、2000 年までの「真」の検索論文を引用した論文のデータは Thomson Reuters 社から購入した。*Publ, Cited, RatePubl, MedCites* の値は、完全計数法と均分式の調節計数法（各著者に論文の著者数の逆数のクレジットを与える）の両方で計算したが、結果に大きな差はなかった（調節計数法の方が、負の 2 項重回帰の適合度がややよかった）ので、以下では調節計数法による結果を示す。

### 4.2.4 負の 2 項重回帰(NBMR)分析

論文の被引用数が 4.2.3 で示した種々の測度にどの程度影響されるかを調べるため、負の 2 項重回帰(NBMR)分析を用いた。目的変数である被引用数は非負の値をとる整数であること、その分布の歪度が極めて大きく、分散が平均に比べて大きいことから、いくつかの研究において NBMR 分析は成功的に用いられている(van Dalen and Henkens, 2001, 2005; Davis et al., 2008; Bornmann and Daniel, 2008; He, 2009; Chen, 2012; Didegah and Thelwall, 2013)。被引用数の対数値（多くは  $\log(C+1)$ ）を目的変数とする線形重回帰もしばしば用いられる(Basu and Lewison, 2005; Figg et al., 2006; Davis and Fromerth, 2007; Davis, 2009)。しかし、NBMR の方がよい結果を与えた (4.4.4 を参照) ので、このモデルを採用することにした。

NBMR 分析では、ケース  $i$  に対する目的変数の値  $y_i$  が、NB 分布

---

<sup>13</sup> 3.3.2 で述べたように、24 誌のうち 6 誌（各分野から 1 誌）の論文については全著者の過去実績データを得た。しかし、これらを用いた分析では明確な結果は得られなかったため、以下では第一著者のデータを用いた結果のみを記す (4.4.2 を参照)。

$$\Pr(y_i = k) = \frac{\Gamma(k + \theta)}{\Gamma(\theta)\Gamma(k + 1)} \left( \frac{\theta}{\mu_i + \theta} \right)^\theta \left( \frac{\mu_i}{\mu_i + \theta} \right)^k \quad [4-1]$$

に従うとする ( $\Gamma(\cdot)$ はガンマ関数)。そしてその期待値  $\mu_i$  を次の回帰式により予測する。

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} \quad [4-2]$$

偏回帰係数  $\beta_0, \beta_1, \beta_2, \dots, \beta_p$  及びパラメータ  $\theta$  の推定値は、 $\{X_{i1}, X_{i2}, \dots, X_{ip}; y_i\}$  の入力データに基づいて算出される。 $\theta$  は  $i$  によらないとする。

分野によって被引用数や説明変数属性の分布がかなり異なることを考慮し、本章では、6つの分野ごとにNBMR分析を行った。しかし、4.1で述べたように、我々はその結果に分野を越えた一般性が見られることを期待する。

目的変数  $y$  は C6 または C11 である。説明変数  $X_1, X_2, \dots, X_p$  は、4.2.3 で示した(a)~(o) であるが、これらの選択が異なる3つの回帰モデル(Model A, Model B, Model C)を設定した。次のように、これらは著者の過去実績に関する属性の選択が異なる。

Model A: *Authors, Insts, Countries, Refs, Price, Figures, Tables, Eqs, Length, Publ, MedCites, Jnt1, Jnt2, Jnt3*

Model B: *Authors, Insts, Countries, Refs, Price, Figures, Tables, Eqs, Length, Cited, RatePubl, Jnt1, Jnt2, Jnt3*

Model C: *Authors, Insts, Countries, Refs, Price, Figures, Tables, Eqs, Length, Age, RatePubl, MedCites, Jnt1, Jnt2, Jnt3*

このように3つのモデルを設定した理由については後述する(4.3.1.3を参照)。

雑誌によって被引用度が異なることを考慮して、NBMRの説明変数に、それぞれの論文が発表された雑誌を表すダミー変数を導入した。各分野ともサンプル論文は4つの雑誌から採られているので、3つのダミー変数  $Jnt1, Jnt2, Jnt3$  を設定する。ベースラインとする1つの雑誌の論文については、 $Jnt1, Jnt2, Jnt3$  の値をすべて0、他の3誌の論文については、当該誌に相当するダミー変数の値を1、その他の変数の値は0とする。各分野で平均被引用数が最も低い雑誌をベースライン誌とする。

NBMRの実行にはSPSS/PASW Version 18のAdvanced Regression Modelを用いた。変数選択は行わず、得られた結果からC6またはC11と有意な関係を持つ変数を識別した。

## 4.3 結果

### 4.3.1 予備的な分析

#### 4.3.1.1 各変数の平均値と標準偏差

#### 第4章 論文の被引用数に対する外在的要因の影響

2つの目的変数（*C6*と*C11*）及び14の説明変数の平均値と標準偏差を表4-1に示す。

*C6*の平均値は分野間で因子4.5くらいの違いがあるが、*C11*の分野間格差は因子2.6に減る。これは、最も被引用数の低い電気分野で*C11/C6*比が3以上になる（他の5分野では2弱）ためである。

説明変数のうち、*Eqs*, *Publ*, *Cited*は特に平均値の分野間格差が大きい。たとえば、物性、電気の2分野では*Eqs*の平均は10-20であるが、消化器分野では用いたサンプル中に数式のある論文は皆無であった。著者の累積業績を示す*Publ*, *Cited*に比べ、効率業績を表す*RatePubl*, *MedCites*の分野間格差はかなり小さい。

以上のように、目的変数、説明変数とも、平均値は分野により相当異なるので、被引用数に影響する要因を検討する研究では対象分野を限定すべきであり、異なる分野の論文が混じった標本を用いることは不適當である。

表4-1 目的変数(*C6*, *C11*)及び14の説明変数の平均値と標準偏差 (括弧内)

分野	物性	無機	電気	生化	生理	消化器
<i>n</i>	230	227	229	240	236	233
<i>C6</i>	10.7 (12.3)	10.7 (11.4)	5.6 (8.6)	21.3 (19.3)	15.8 (14.8)	25.0 (32.0)
<i>C11</i>	19.5 (24.0)	19.0 (22.4)	17.3 (30.9)	35.0 (33.2)	26.4 (25.3)	45.8 (67.6)
<i>Authors</i>	3.37 (2.06)	4.27 (1.97)	2.68 (1.52)	4.88 (2.13)	4.14 (2.16)	6.61 (2.99)
<i>Insts</i>	1.86 (0.86)	1.70 (0.80)	1.44 (0.71)	1.81 (0.85)	1.55 (0.73)	1.67 (0.95)
<i>Countries</i>	1.39 (0.56)	1.29 (0.55)	1.17 (0.43)	1.28 (0.54)	1.18 (0.44)	1.15 (0.42)
<i>Refs</i>	27.3 (13.1)	34.3 (18.6)	18.9 (17.2)	41.7 (16.1)	36.7 (15.4)	31.0 (14.3)
<i>Price</i>	32.4 (19.6)	26.6 (15.8)	33.0 (21.0)	40.1 (18.8)	36.6 (18.1)	33.9 (20.4)
<i>Figures</i>	5.89 (3.79)	4.83 (2.88)	8.63 (4.88)	6.49 (2.31)	6.66 (3.24)	3.55 (2.39)
<i>Tables</i>	0.87 (1.56)	2.98 (1.88)	1.18 (1.59)	1.45 (1.45)	0.91 (1.33)	1.74 (1.58)
<i>Eqs</i>	13.27 (17.46)	2.13 (4.22)	19.25 (20.53)	0.80 (3.13)	1.24 (3.93)	0.00 (0.00)
<i>Length</i>	7.28 (3.18)	7.10 (2.72)	8.28 (3.43)	10.70 (3.40)	9.94 (4.11)	6.20 (2.16)
<i>Publ</i>	6.41 (12.68)	10.74 (29.01)	3.58 (5.17)	2.54 (4.66)	3.48 (5.39)	7.69 (19.04)
<i>Cited</i>	35.2 (226.6)	98.1 (492.3)	6.1 (17.7)	26.6 (75.3)	24.7 (60.6)	53.9 (315.1)
<i>Age</i>	7.09 (6.76)	8.73 (8.40)	4.61 (5.93)	5.41 (5.94)	5.90 (6.77)	8.36 (7.43)
<i>RatePubl</i>	0.85 (1.13)	0.92 (1.40)	0.69 (0.96)	0.38 (0.40)	0.57 (0.55)	0.86 (1.09)
<i>MedCites</i>	0.074 (0.197)	0.072 (0.119)	0.046 (0.182)	0.171 (0.259)	0.122 (0.269)	0.049 (0.134)

#### 4.3.1.2 目的変数と説明変数の間の相関

NBMRを行う前に、*C6*及び*C11*と個々の説明変数の間の相関を求めた。*C6*と各説明変数の間のSpearman順位相関係数( $\rho$ )を表4-2に示す。この表には、相関が有意( $p < 0.05$ )である場合のみ相関係数の値を示している。*C11*についても類似の結果であった。

*Refs*, *Price*, *Length*がすべて、あるいはほとんどの分野で有意な正の相関を示し、*Insts*, *Figures*, *Cited*, *RatePubl*, *MedCites*も半数以上の分野で有意な正の相関を示す (*Age*も3分野で有意であるが、 $\rho$ の値が正の場合と負の場合があるので、明確な傾向を持つとは言えない)。全体を通じて $\rho$ の絶対値はほとんど0.4以下、最大でも0.6で、著しく強い相関は見られない。

表 4-2 C6 と各説明変数の間の Spearman 順位相関関数( $\rho$ )

(相関が 5%有意である  $\rho$  の値のみを示す)

説明変数	物性	無機	電気	生化	生理	消化器
<i>Authors</i>		0.204 **				
<i>Insts</i>		0.199 **	0.171 **			0.246 **
<i>Countries</i>			0.134 *			0.283 **
<i>Refs</i>	0.254 **	0.395 **		0.312 **	0.494 **	0.382 **
<i>Price</i>	0.376 **	0.357 **	0.188 **	0.555 **	0.488 **	0.392 **
<i>Figures</i>		0.375 **		0.153 *	0.442 **	0.132 *
<i>Tables</i>						0.153 *
<i>Eqs</i>						
<i>Length</i>	0.138 *	0.439 **	0.130 *	0.363 **	0.601 **	0.349 **
<i>Publ</i>			0.215 **			0.239 **
<i>Cited</i>	0.190 **		0.277 **			0.242 **
<i>Age</i>			0.190 **	-0.169 **	-0.225 **	
<i>RatePubl</i>			0.191 **		0.166 *	0.285 **
<i>MedCites</i>	0.219 **		0.196 **	0.178 **		

\*\* 1%有意 \* 5%有意

#### 4.3.1.3 説明変数間の相関

表 4-3 に、説明変数のそれぞれの組み合わせにどの程度相関が見られるかを示した。この表には、Spearman 順位相関が有意( $p < 0.05$ )な分野の数を示す。表中に枠で囲んだ、著者の過去業績を示す 5 つの変数グループ{*Publ*, *Cited*, *Age*, *RatePubl*, *MedCites*}、研究協力の程度を示す 3 つの変数グループ{*Authors*, *Insts*, *Countries*}、及び論文の可視性を示す 3 つの変数グループ{*Refs*, *Figures*, *Length*}の中では、それぞれ明確な正の相関がある。但し、最初のグループの中で、2 つの効率業績指標、*RatePubl* と *MedCites* の間には、有意な相関を示す分野は多くない。*Length* は *Tables*, *Eqs* と半分以上の分野で正の相関を持つ。一方、*Price* は *Tables*, *Eqs*, *Age* と負の相関を示す傾向がある。

4 分野以上で有意な相関を示す変数の組について、その詳細を表 4-4 に示す。

著者の過去業績を示す変数グループの中で、{*Publ*, *Cited*, *Age*}の相互間の相関は特に強く、ほとんどの  $\rho$  が 0.7 を越える。しかし、これら 3 つの累積業績変数と 2 つの効率業績変数(*RatePubl*, *MedCites*)の間の相関は、*Publ*-*RatePubl* を除いてそれほど強くはない。{*Authors*, *Insts*, *Countries*}の相互間、及び{*Refs*, *Figures*, *Length*}相互間には、やや強いあるいは中程度の相関( $\rho = 0.5 \sim 0.7$ )を示す例が多いが、非常に強い相関は見られない。

このように説明変数の間にかなりの相関関係が存在するので、表 4-2 で示した目的変数と有意な相関を持つ変数が、重回帰分析で有意になるとは限らない。たとえば、*Refs* と *Length* はどちらも全分野で C6 及び C11 と有意な相関を示すが、両変数の間にもかなり強い相関があるので、重回帰分析ではこれらのどちらか（またはどちらも）有効な予測変数にならないかもしれない。

表 4-3 説明変数間の相関

数字は、Spearman 順位相関係数( $\rho$ )が有意( $p < 0.05$ )な分野の数。括弧内数字は、そのうち  $\rho$  の値が負である分野の数。

	Authors	Insts	Countries	Refs	Figures	Length	Price	Tables	Eqs	Publ	Cited	Age	RatePubl
Insts	6												
Countries	6	6											
Refs	1	3	2										
Figures	3 (1)	3 (2)		5									
Length	3 (1)	3 (1)	1	6	6								
Price	1	2	1	2	1	2							
Tables	3		1	1	3 (1)	3	3 (3)						
Eqs	2 (2)		1	3	2	4	2 (2)	3 (1)					
Publ	1 (1)	2					1	1 (1)					
Cited	1	3					2	1		6			
Age	3	2				1 (1)	3 (3)	1		6	6		
RatePubl		1		1	1	1	2		1	6	6	5	
MedCites		1				1	4 (1)	1 (1)		5	6	6	3

表 4-4 4分野以上で有意な相関を示す説明変数の組に対する Spearman 順位相関係数

説明変数の組	物性	無機	電気	生化	生理	消化器
(Authors, Insts)	0.555 **	0.550 **	0.441 **	0.474 **	0.436 **	0.169 **
(Authors, Countries)	0.376 **	0.324 **	0.243 **	0.274 **	0.281 **	0.134 *
(Insts, Countries)	0.647 **	0.627 **	0.613 **	0.498 **	0.541 **	0.487 **
(Refs, Figures)	0.113	0.412 **	0.196 **	0.342 **	0.497 **	0.287 **
(Refs, Length)	0.548 **	0.634 **	0.478 **	0.614 **	0.711 **	0.705 **
(Figures, Length)	0.494 **	0.693 **	0.486 **	0.660 **	0.763 **	0.598 **
(Eqs, Length)	0.421 **	0.099	0.352 **	0.160 *	0.376 **	- <sup>a</sup>
(Price, MedCites)	0.273 **	-0.169 *	0.122	0.196 **	0.069	0.146 *
(Publ, Cited)	0.821 **	0.862 **	0.731 **	0.811 **	0.818 **	0.845 **
(Publ, Age)	0.731 **	0.779 **	0.816 **	0.762 **	0.749 **	0.638 **
(Publ, RatePubl)	0.687 **	0.785 **	0.835 **	0.706 **	0.706 **	0.743 **
(Publ, MedCites)	0.308 **	0.321 **	0.126	0.328 **	0.238 **	0.167 *
(Cited, Age)	0.747 **	0.841 **	0.817 **	0.825 **	0.756 **	0.735 **
(Cited, RatePubl)	0.442 **	0.553 **	0.470 **	0.495 **	0.467 **	0.473 **
(Cited, MedCites)	0.622 **	0.625 **	0.538 **	0.691 **	0.570 **	0.473 **
(Age, RatePubl)	0.180 **	0.335 **	0.544 **	0.334 **	0.259 **	0.073
(Age, MedCites)	0.310 **	0.461 **	0.383 **	0.435 **	0.283 **	0.262 **

\*\* 1%有意 \* 5%有意

a) 消化器分野では、全サンプル論文について *Eqs* の値が 0。

説明変数間に非常に強い相関があると、多重共線性により重回帰分析の回帰式が不安定になるので、そのような回帰モデルを用いることは避けるべきである。上に述べた結果から、多くの分野で  $\rho > 0.7$  となる変数組み合わせは、同じモデルの説明変数に含めないことにした。具体的には次の通りである。

- *Publ, Cited, Age* の 2 つ以上を 1 つのモデルに含めない。
- *Publ* と *RatePubl* も 1 つのモデルに含めない。

この考えにより、4.2.4 で述べた 3 つの回帰モデル(Model A, Model B, Model C)を設定した。なお、上記の 2 条件からは Model B に *MedCites* を含めることは排除されないが、*Cited* と似た性格の指標なので含めなかった。グループ{*Authors, Insts, Countries*}内の相互間、及びグループ{*Refs, Figures, Length*}内の相互間では  $\rho$  が 0.7 を越すことはないで、これらを分離することはしなかった。

### 4.3.2 NBMR の結果

#### 4.3.2.1 全体的な適合度(Goodness of fit)ーモデル間の比較

NBMR の全体的適合度を示す指標にはいろいろのものがある(Long, 1997, pp. 85-113)。表 4-5 では、これらのうち Akaike 情報量規準(*AIC*)と調整済み擬決定係数(adjusted pseudo coefficient of determination; pseudo  $R^2$ )を、3 つのモデル間で比較した。

*AIC* は小さいほど、pseudo  $R^2$  は大きいほど (1 に近いほど) 当てはまりがよい。*AIC*、pseudo  $R^2$  とも、モデル間の優劣は分野により異なるが、どの分野においてもモデル間に著しい差はない。強いて言えば、Model C が他の 2 つに比べて分野を通しての安定性が高い。

い。

表 4-5 3つの NMBR モデルの適合度比較

(a) Akaike 情報量規準 (AIC)

分野	目的変数: C6			目的変数: C11		
	モデルA	モデルB	モデルC	モデルA	モデルB	モデルC
物性	1511.3	1511.2	1515.6	1783.5	1782.5	1787.1
無機	1466.4	1475.9	1470.9	1721.7	1728.3	1725.8
電気	1214.0	1211.4	1211.9	1687.2	1685.5	1686.8
生化	1811.4	1814.9	1814.1	2054.3	2059.1	2056.8
生理	1613.9	1613.2	1612.1	1872.4	1871.1	1868.2
消化器	1814.7	1811.0	1813.4	2088.1	2084.9	2087.5

(b) 調整済み擬決定係数(pseudo  $R_e^2$ )

分野	目的変数: C6			目的変数: C11		
	モデルA	モデルB	モデルC	モデルA	モデルB	モデルC
物性	0.261	0.261	0.251	0.236	0.239	0.228
無機	0.332	0.304	0.322	0.318	0.299	0.310
電気	0.276	0.285	0.287	0.283	0.288	0.288
生化	0.441	0.433	0.438	0.424	0.413	0.421
生理	0.531	0.533	0.537	0.483	0.485	0.493
消化器	0.487	0.495	0.492	0.488	0.495	0.492

#### 4.3.2.2 被引用数への影響要因—モデル間の比較

3つの回帰モデルにおいて、各説明変数に対する偏回帰係数が有意となった分野の数を表4-6に示す。ここでは、有意確率  $p$  が 0.1 未満の場合を有意とした。

最も明瞭な効果を示すのは *Price* で、どの分野でも、それが高い論文ほど被引用数が高いと予測される。*Refs*, *Authors*, *Figures* も、いくつかの分野で被引用数に対して正の影響を与える。これらの傾向はモデルに依らない。

モデルにより選択を行った著者実績属性変数については、累積業績指標の *Publ*, *Cited*, *Age* は、ほとんどの場合被引用数に影響を与えないという結果であった。一方効率業績指標の *RatePubl* と *MedCites* は、いくつかの分野で正の影響を与えた。従って、3つのモデルの中では、この2つの変数を含む Model C が好ましいと考えられる。

#### 4.3.2.3 被引用数の予測要因—Model C に対する結果

以上の結果より、適合度、選択される説明変数のどちらの観点からも、Model C が比較的適切であると考えられる。そこで、以下では Model C を用いた結果を示す。



表 4-6 各説明変数の被引用数予測有意性

数字は、各変数に対する偏回帰係数が有意( $p < 0.1$ )な分野の数。括弧内数字は、そのうち偏回帰係数の値が負の分野の数。

説明変数	目的変数: C6			目的変数: C11		
	モデルA	モデルB	モデルC	モデルA	モデルB	モデルC
<i>Authors</i>	2	3	2	3(1)	2	2
<i>Insts</i>		1(1)	2(2)	1	1	2(1)
<i>Countries</i>	1	1	1	1	1	1
<i>Refs</i>	3	3	3	3	3	3
<i>Price</i>	6	6	6	6	6	6
<i>Figures</i>	1	1	1	2	2	2
<i>Tables</i>	1	1	1	1	1	1
<i>Eqs</i>	1(1)	1(1)	1(1)	1(1)	2(2)	1(1)
<i>Length</i>	1	1	1	2(1)	2(1)	2(1)
<i>Publ</i>	2(1)	-	-	2(1)	-	-
<i>Cited</i>	-		-	-		-
<i>Age</i>	-	-		-	-	1(1)
<i>RatePubl</i>	-	2	4(1)	-	2	3(1)
<i>MedCites</i>	1	-	2	2	-	2

6つの分野に対するNBMRの結果を表4-7に示した。この表に示す各説明変数に対する偏回帰係数は、それぞれの変数の目的変数に対する影響の相対的強さを比較できるように、回帰係数 $\beta_j$ に変数の標準偏差 $s_j$ を乗じた $x$ 標準化回帰係数を示している。表4-7には負の2項分布パラメータ $\theta$ に対する推定値も記載している。なお、消化器分野ではすべての論文において*Eqs*(数式の数)が0であったので、説明変数に含まれない。

C6とC11に対する回帰の結果はほぼ同様であることが表4-7から判る。

有意性の高さ、 $x$ 標準化回帰係数の大きさのいずれからも、*Price*が最も重要な被引用数への影響要因である。次いで*Refs*が半分の分野で、強い、あるいは中程度の有意な正の予測因子である。*Authors*, *Figures*及び効率業績指標である*RatePubl*, *MedCites*も、正の予測因子となる複数の分野があるが、*RatePubl*は無機分野でC6とC11に対しては負の予測因子となる。無機以外の5分野では、論文が発表される雑誌により、予測される被引用数にかなり明確な差がある。

上記以外の変数には、被引用数に対する明確な影響は認められない。*Insts*, *Length*は複数の分野で有意であるが、回帰係数の符号が一定していない。

表4-8は、本章での研究で得られた各変数の被引用数予測力を、多くの変数を統合的に検討した表2-7に示す研究の結果と比較したものである。

表4-7 モデルCに対するNBMRの $x$ 標準化回帰係数 ( $s_j\beta_j$ )

(a) 目的変数:  $C6$

	物性	無機	電気	生化	生理	消化器
$n$	230	227	229	240	236	233
$\theta$	1.44	2.15	1.25	2.95	3.11	1.90
$s_j\beta_j$						
<i>Authors</i>	0.174 *	0.052	-0.115	0.116 *	-0.012	0.096
<i>Insts</i>	-0.161 +	0.119	0.166	0.003	0.083	-0.107 +
<i>Countries</i>	-0.074	-0.099	-0.040	-0.030	-0.053	0.107 +
<i>Refs</i>	0.273 **	0.062	0.190 *	0.122 *	0.089	0.103
<i>Price</i>	0.331 **	0.233 **	0.250 **	0.391 **	0.354 **	0.251 **
<i>Figures</i>	-0.048	0.083	0.036	-0.018	0.141 +	0.099
<i>Tables</i>	-0.007	-0.085	-0.085	0.015	-0.050	0.281 **
<i>Eqs</i>	-0.033	-0.129 *	-0.025	0.009	-0.042	-
<i>Length</i>	0.079	0.239 *	0.124	0.094	0.134	-0.150
<i>Age</i>	0.056	-0.035	0.008	-0.009	-0.071	-0.005
<i>RatePubl</i>	0.099 +	-0.112 +	0.194 *	0.025	0.058	0.118 *
<i>MedCites</i>	-0.002	0.144 **	0.102	0.071 +	0.041	-0.022
<i>Dummy1</i>	0.254 **	0.122	0.549 **	0.123 *	0.453 **	0.333 **
<i>Dummy2</i>	0.049	0.000	0.628 **	0.274 **	0.462 **	0.792 **
<i>Dummy3</i>	0.338 **	0.028	0.101	0.303 **	0.264 **	0.627 **

\*\* 1%有意 \* 5%有意 + 10%有意

(b) 目的変数:  $C11$

	物性	無機	電気	生化	生理	消化器
$n$	230	227	229	240	236	233
$\theta$	1.27	1.86	1.04	2.58	2.44	1.69
$s_j\beta_j$						
<i>Authors</i>	0.175 *	0.052	-0.143	0.136 **	-0.019	0.095
<i>Insts</i>	-0.169 +	0.085	0.175	-0.006	0.113 +	-0.098
<i>Countries</i>	-0.094	-0.092	-0.066	-0.021	-0.043	0.162 *
<i>Refs</i>	0.213 **	0.046	0.343 **	0.123 *	0.102	0.099
<i>Price</i>	0.313 **	0.200 **	0.225 **	0.393 **	0.314 **	0.199 **
<i>Figures</i>	-0.065	0.108	0.095	-0.034	0.139 +	0.165 +
<i>Tables</i>	0.022	-0.012	0.008	0.044	-0.039	0.362 **
<i>Eqs</i>	0.000	-0.130 *	-0.074	0.032	-0.079	-
<i>Length</i>	0.123	0.256 *	0.137	0.085	0.148	-0.236 *
<i>Age</i>	0.053	-0.020	0.041	0.024	-0.107 *	-0.018
<i>RatePubl</i>	0.115 +	-0.125 *	0.164 *	0.010	0.073	0.099
<i>MedCites</i>	-0.012	0.126 *	0.066	0.081 +	0.026	-0.025
<i>Dummy1</i>	0.248 **	0.096	0.577 **	0.125 *	0.407 **	0.335 **
<i>Dummy2</i>	0.040	0.006	0.534 **	0.270 **	0.411 **	0.814 **
<i>Dummy3</i>	0.325 **	-0.004	-0.002	0.315 **	0.205 **	0.648 **

\*\* 1%有意 \* 5%有意 + 10%有意

第4章 論文の被引用数に対する外在的要因の影響

表 4-8 論文被引用数予測力に関する種々の研究の比較(本章での研究を含む)

文献	対象分野	サンプルサイズ (n)	要因を表す測度 <sup>a)</sup>														
			著者数	所属機関 数 <sup>b)</sup>	所属国 数 <sup>c)</sup>	著者の生 産性	著者の引 用影響度	著者の経 歴・活動期 間	雑誌のイ ンパクト	参考文献 数	参考文献 の最新性	論文長	図の数	表の数	数式の数		
Bornmann & Daniel (2008)	化学	1,837	B					A									C
Chen (2012)	4つのトピック	1,300 - 6,800 for a topic	A									A					C
Davis et al. (2008)	生理学	1,619	A								A	A					C
Didegah & Thelwall (2013)	ナノ科学・技術	50,162	C	A	C						A	A					
Fu & Aliferis (2010)	内科学	3,788	C	C		C		A			A						
Haslam et al. (2008)	社会心理学・人格心理学	308	C			A						A	A	A	C	C	
He (2009)	生医学	1,860	A		A			B		C		A					C
Lokker et al. (2008)	臨床医学	1,261	A									A					C
Peng & Zhu (2012)	インターネット研究	18,580	A							A	A						A
Peters & van Raan (1994)	化学工学	226	B				A				A	A	A				B
Stewart (1983)	地球科学	139	C						A	C		A	B				A
Van Dalen & Henkens (2001)	人口学	1,371	B								A						A
Walters (2006)	犯罪心理学	428	B								A	B					C
This research	6つの分野	227 - 240 for a field	B	C	C	B	B	C	A	A	A	A	A	C	C	C	C

a) A:強いまたは明確な影響力 B:弱い影響力またはモデルにより予測力変化 C:有意性なしまたは負の影響力

b) 国内機関共著の有無、他の国内機関の共著者数を含む。

c) 国際共著の有無、他国の共著者数を含む。

4.3.2.4 個々の論文に対する被引用数予測の精度

図4-1に、生理分野及び物性分野における論文*i*に対する*C6*の観測値( $C6_i$ )と、Model CによるNBMRから得られたその予測期待値 $\mu_i$ の関係を示す。これらの分野はそれぞれ、pseudo  $R^2$ が6分野中最高と最低の分野である(表4-5bを見よ)。 $\mu_i$ が大きくなるほど残差が大きくなる傾向が見られる。NBMRにおいては、式[4-2]に示すように予測される値は $\ln(\mu)$ なので、残差( $C6_i - \mu_i$ )は概ね $\mu_i$ に比例して大きくなると思われる。そこで、これらの2分野について、 $(C6_i - \mu_i)/\mu_i$ (この量を相対残差と呼ぶ)と $\mu_i$ の関係を図4-2に示した。相対残差は $\mu_i$ にあまり依存しないことが判る。

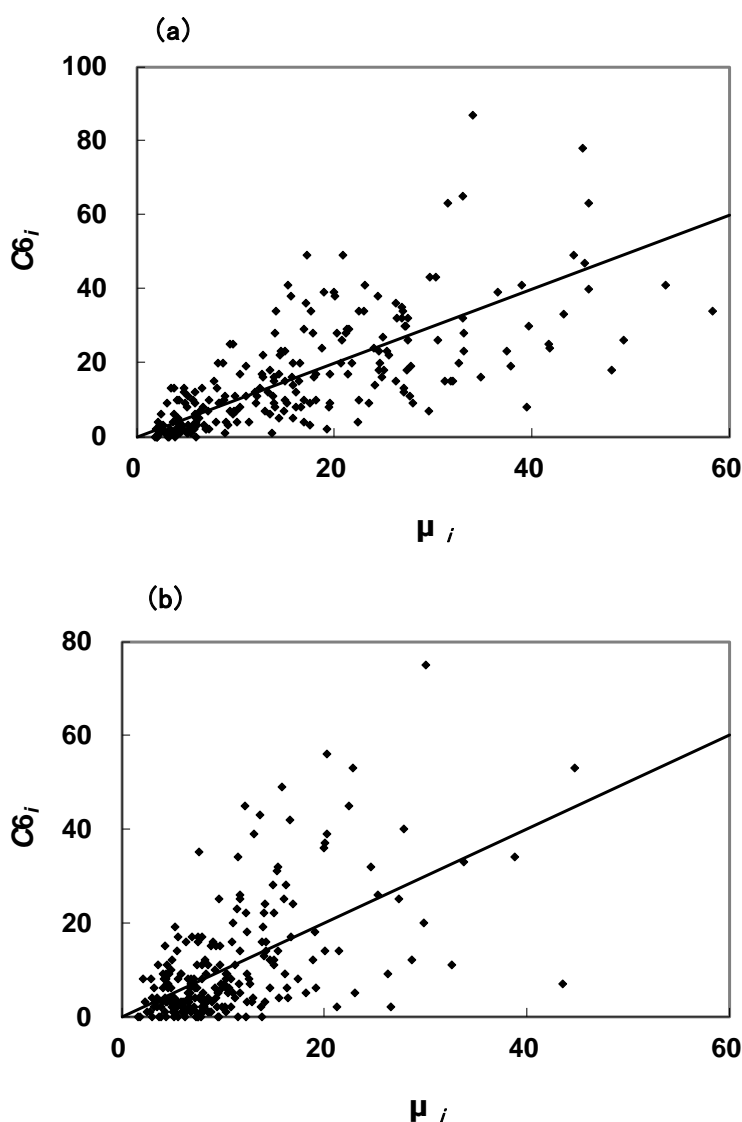


図4-1 各論文に対する*C6*の観測値( $C6_i$ )とNBMRによる予測期待値( $\mu_i$ )の関係

(a) 生理分野 (b) 物性分野

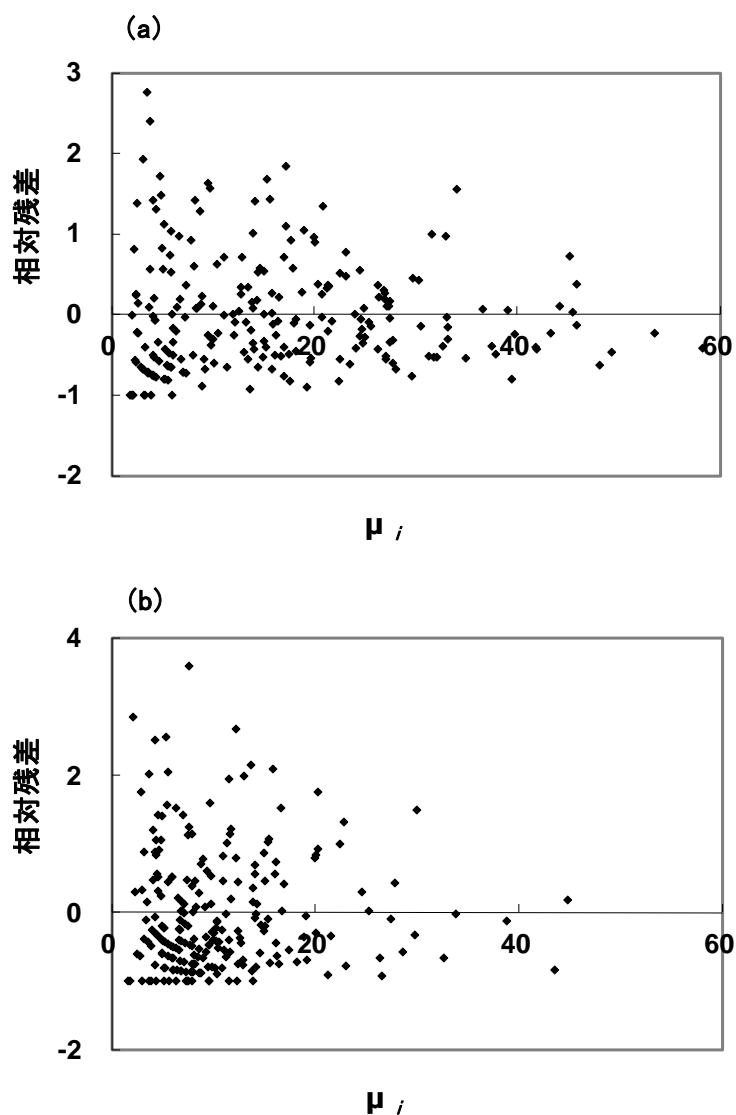


図 4-2 各論文に対する  $C6$  の予測期待値( $\mu_i$ )と相対残差の関係

(a) 生理分野 (b) 物性分野

図 4-2 から、相当数の論文で、被引用数の観測値がその期待値の 2 倍以上、すなわち  $(C6_i - \mu_i) / \mu_i > 1$  であることが判る。これを見ると予測精度はよくないようであるが、 $\mu_i$  は  $C6_i$  の直接の予測値ではなく、式[4-1]と式[4-2]から判るように、 $C6_i$  の負の 2 項分布の期待値であることに注意する必要がある。

NBMR から予測される  $C6$  及び  $C11$  の分布は次の手続きにより得られる。

- (a) それぞれの論文  $i$  に対して式[4-1]から予測される被引用数の確率分布  $\Pr(C6_i = k)$  及び  $\Pr(C11_i = k)$  を計算する。

(b) これらの分布を全論文について積算する。

この予測分布と実際の分布を比較したものを図4-3に示す。ここに示すのは無機分野の例で、この分野の適合度は6分野中で中程度である。実測値の分布は変動が大きいですが、NBMRによる予測はそれを平滑化した曲線をよく再現している。

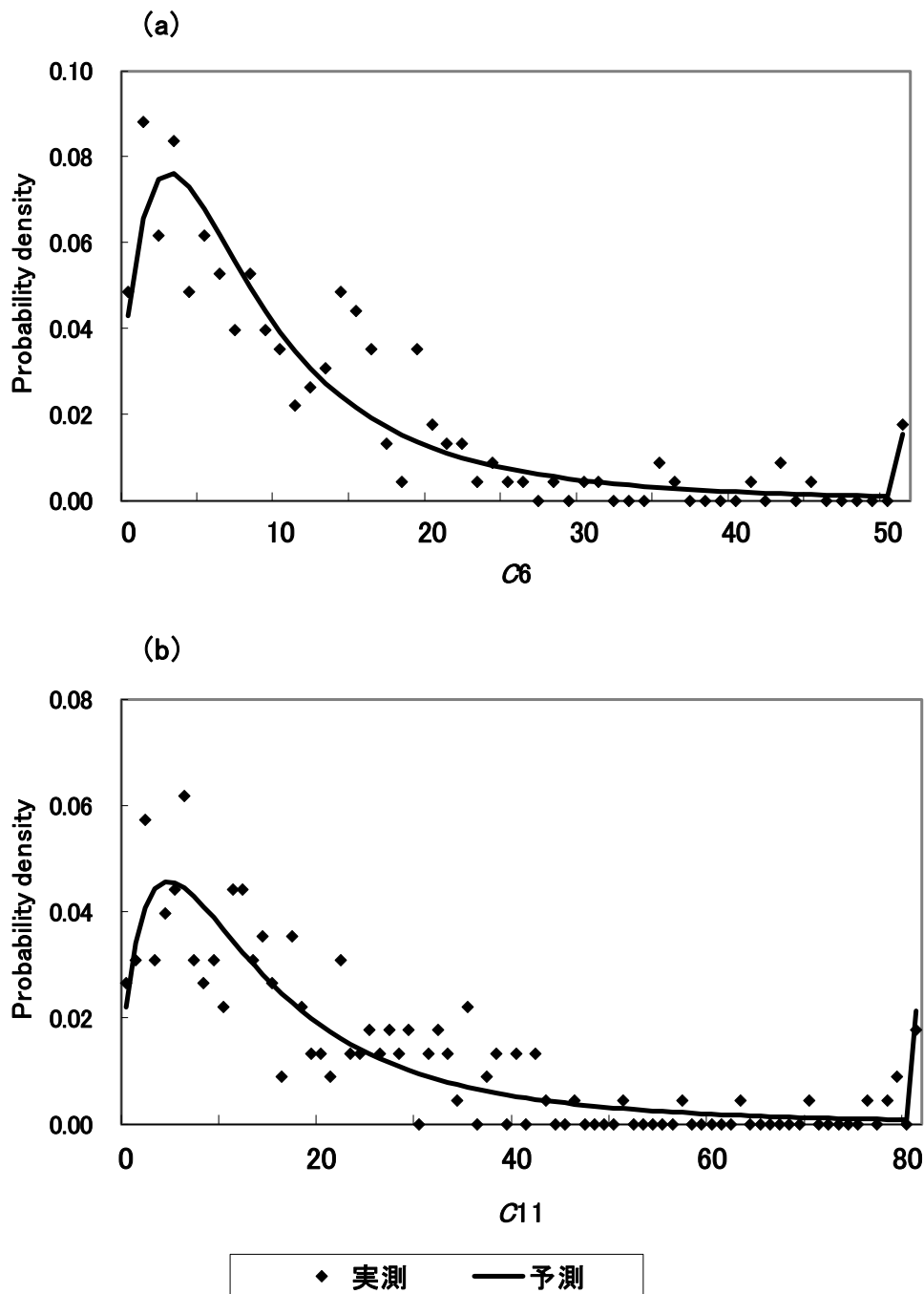


図4-3 NBMRの予測による被引用数確率分布と実際の分布との比較(無機分野の場合)

(a) C6の分布 (b) C11の分布

## 4.4 考察

### 4.4.1 被引用数に対する重要な影響要因

6つの異なる分野それぞれに対し、論文発表後6-7年間及び11-12年間の被引用数(*C6*, *C11*)を、3~5個の有意な説明変数により予測することができた。NBMRに対するpseudo  $R^2$ は、分野により*C6*では0.25~0.54、*C11*では0.23~0.50であった。有意な説明変数となる属性は、分野を越えてある程度の共通性があり、*C6*と*C11*でほぼ同様であった。

Price 指数(*Price*)が全ての分野で最も強力な被引用数への影響属性であることが見出された。この種の属性(参考文献の最新性を示す尺度)に注目した研究はいくつかある(Stewart, 1983; Peters and van Raan, 1994; Haslam et al., 2008)。これらの研究では、参考文献の最新性は論文の被引用数と正の相関を持つことが示されているが、いずれも単一の分野の比較的少数のサンプル論文( $n < \sim 300$ )を用いている(表4-8参照)。Price 指数が被引用数の予測変数として極めて重要であることを全分野で示したことは、本章で述べた研究の最も主要な発見の一つである。

次に重要な説明変数は参考文献数(*Refs*)であったが、これは多くの先行研究で被引用数との有意な関係が示されている(2.1.2.4及び表2-5を参照)。

著者数の多い論文はよく引用されるという報告は多いが、体系的な重回帰分析ではその主張は必ずしも支持されていない(2.1.2.1及び表2-2を参照)。NBMRを用いた本章での研究でも、著者数(*Authors*)は6分野中2分野で(中程度に)有意な予測変数となるに留まった。著者数は被引用数に対しそれほど強い影響要因ではないことが示唆される。Bornmann and Daniel (2008)は、著者数と被引用数の間の相関は、被引用期間が長くなるとなくなる傾向にあると報告しているが、ここで6年及び11年という比較的長い被引用ウィンドウを用いたことが、これと関係がある可能性がある。

著者実績変数が被引用数に及ぼす影響については、次の細節で考察する。

### 4.4.2 引用への著者のハロー効果はあるか?

2.1.2.2で述べたように、過去実績(発表論文数とそれら論文が得た被引用数)に富む著者の論文はその後多くの引用を得る傾向があるとする報告がいくつかある。本章では、5つの著者実績指標が論文の被引用数に及ぼす効果を検討した(いずれも第一著者に対するもの)。その結果、3つの累積実績指標(*Publ*, *Cited*, *Age*)は、いずれも被引用数に対する効果がほとんど認められなかった。これに対し、2つの効率実績指標(*RatePubl*, *MedCites*)はいくつかの分野で有意な影響を示したが、それほど明確なものではなかった。

## 第4章 論文の被引用数に対する外在的要因の影響

効率実績指標の方が累積実績指標よりも被引用数のよい予測変数であるという結果は、Danell (2011)及び Hönekopp and Khan (2012)の研究と一致する。しかし、我々の分析でその効果がそれほど強く現れなかったのは、我々が第一著者のデータを用いたことによるのかもしれない。Danell は最も得点の高い著者のデータを用い、Hönekopp and Khan は単独著者論文のみをサンプルに選んでいる。第一著者のデータのみを用いたことは、引用に対する著者のハロー効果を弱めた可能性があり、本章での研究の限界の一つと考えられる。我々は、各分野から1誌ずつを選んで全著者の過去実績を調べ、最高得点著者の実績を用いた回帰を試みたが、明確な結果は得られなかった。これはサンプルが小さ過ぎた( $n \approx 50-60$ )ことによるのかもしれない。

### 4.4.3 説明変数間の相互作用

表4-2と表4-7を比較すると、被引用数と有意な相関がある説明変数が、必ずしもNBMRで有意な予測変数となるとは限らない。その典型的な変数は *Length* で、ほとんど全ての分野で被引用数と正の有意な相関があるが、NBMR で有意な予測変数となる場合は著しく減る。表4-3及び表4-4に見るように、*Length*は多くの分野で *Refs*, *Figures*, *Tables*, *Eqs* などの変数と正の相関があり、NBMR ではこれらの変数が *Length* より優先的に選択されたと考えられる。

### 4.4.4 線形重回帰(LMR)分析と NBMR の適合度比較

本章での被引用数予測には NBMR を用いたが、目的変数に  $\log(C+1)$  を用いる LMR もよく使われる。しかし、この両者に共通に適用できる適合度尺度はあまりないので、同一標本に対する2つの回帰の適合度を比較することは難しい。LMR に通常使われる分散比 ( $F$ ) は NBMR には使われない。LMR における決定係数( $R^2$ )と NBMR における pseudo  $R^2$  はやや意味が異なる。 $AIC$  は、LMR にも NBMR にも使われるが、それぞれの値を単純に比較するには問題がある。そこで、次の2つの尺度により、同じデータに適用した NBMR の結果と LMR の結果を比較した。

#### (a) 平均平方相対残差(mean squared relative residuals)

平均平方相対残差  $MSRR$  は、4.3.2.4 で導入した相対残差の平方平均である。メンバー  $i$  に対する目的変数の観測値を  $y_i$ 、予測値(期待値)を  $\mu_i$  とすると、 $MSRR$  は次式で与えられる。

$$MSRR = \sum_{i=1}^n \left( \frac{y_i - \mu_i}{\mu_i} \right)^2 / n \quad [4-3]$$

4.3.2.4 で述べたように、NBMR においては、残差( $y_i - \mu_i$ )は概ね  $\mu_i$  に比例して大き



くなる。log  $y_i$  を従属変数とする LMR でも、 $y_i$  の残差は期待値に比例して増加するので、両者の性能を *MSRR* によって比較することは合理的である。

(b) 適合度のカイ二乗値

NBMR、LMR それぞれの結果から、被引用数に対する予測度数分布が得られる (NBMR における予測分布を得る方法は 4.3.2.4 で述べた)。予測分布による期待値の度数がほぼ均等になるように、被引用数を  $m$  の区間に分ける (ここでは  $m=10$  とする)。区間  $i (1 \leq i \leq m)$  に入る期待(予測)度数を  $E_i$ 、観測度数を  $O_i$  とすると、適合度カイ二乗の値( $\chi^2$ )は次式で与えられる。

$$\chi^2 = \sum_{i=1}^m \frac{(O_i - E_i)^2}{E_i} \quad [4-4]$$

Model C で選択した説明変数の組を使った NBMR と LMR による *MSRR* の比較を表 4-9 に示す。NBMR に対する *MSRR* は、対応する LMR のその 1/2~1/5 である。

表 4-9 平均平方相対残差(*MSRR*) による NBMR と LMR の適合度比較

分野	NBMR		LMR	
	C6	C11	C6	C11
物性	0.78	0.86	2.34	2.72
無機	0.60	0.64	1.27	1.43
電気	1.37	1.29	6.36	6.78
生化	0.44	0.47	0.73	0.81
生理	0.49	0.54	0.96	1.19
消化器	0.64	0.62	1.56	1.48

また、 $\chi^2$  に対する同様な比較を表 4-10 に示す。NBMR における  $\chi^2$  は、唯一の場合 (電気分野の C11) を除き、予測分布が観測分布と異ならないという帰無仮説を棄却しない (有意確率  $p > 0.05$ ) のに対し、LMR では全ての場合にその帰無仮説は強く棄却される ( $p < 0.001$ )。

これらのことから、NBMR の方がはるかに実分布との当てはまりがよいと言える。

図 4-4 に、無機分野における LMR による予測分布と観測分布を比較した (この分野は、NBMR、LMR とともに観測分布に対する予測分布の当てはまりのよさが中程度)。この図と NBMR に対する図 4-3 と比べても、NBMR の予測が優れていることが明らかである。特に低 C 領域において NBMR の当てはまりがよいことが判る。

表 4-10 カイ二乗値( $\chi^2$ ) とその有意確率( $p$ ) による NBMR と LMR の適合度比較

目的変数	分野	NBMR		LMR	
		$\chi^2$	$p$	$\chi^2$	$p$
C6	物性	3.73	0.928	188.5	0.000
	無機	11.56	0.239	139.0	0.000
	電気	6.34	0.501	259.5	0.000
	生化	7.82	0.553	55.7	0.000
	生理	12.82	0.171	43.9	0.000
	消化器	3.51	0.941	89.7	0.000
C11	物性	5.52	0.786	194.7	0.000
	無機	10.31	0.326	146.9	0.000
	電気	17.75	0.038	143.4	0.000
	生化	6.38	0.701	65.9	0.000
	生理	13.12	0.157	48.1	0.000
	消化器	3.73	0.928	64.9	0.000

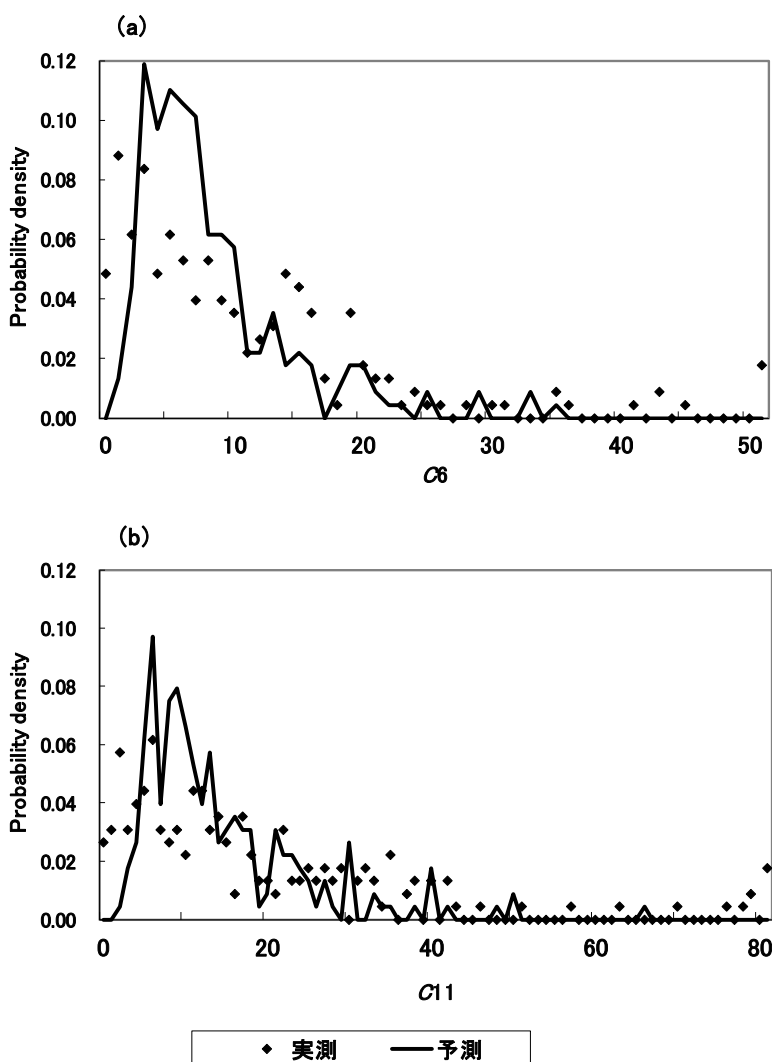


図 4-4 LMR の予測による被引用数確率分布と実際の分布との比較(無機分野の場合)

(a) C6 の分布 (b) C11 の分布

#### 4.4.5 サンプルングの問題

4.2.1 で述べたように、本章の研究で用いたサンプルは、各分野で選んだ4つの雑誌から、それぞれ50~60の論文(“articles”)を無作為抽出したものである。これに関して2つの問題が指摘されるかもしれない。

その一つは、サンプルサイズが比較的小さいことである(各分野について  $n=230-240$ )。これは、データを得るのに労力を要する説明変数を多く用いたことによる。*Authors, Insts, Countries, Refs* は、用いたデータ源(WoS)から比較的簡単に得られるが、*Price* は各サンプル論文の個々の参考文献の発行年を調べなければならない。*Figures, Tables, Eqs* は、個々のサンプル論文に直接当たってカウントした。*Length* については単純なページ数ではなく、各雑誌のページあたり文字数により規格化した。5つの著者実績変数(*Publ, Cited, Age, RatePubl, MedCites*)のデータを得るための、著者の過去発表論文の検索と同定、及びそれらの論文がサンプル論文の発表年(2000)までに得た被引用数の測定には、3.3 で述べたように、最も大きな労力が必要であった。

より大きなサンプルを用いれば、有意な説明変数が増える可能性があったことは否定できない。著者実績変数も、より明確な予測変数になったかもしれない。しかし、逆にサンプルが大きすぎると、被引用数の予測にあまり重要でない変数まで影響を持つという結果が得られてしまうこともある。本章の研究で有意であった変数は、確実に被引用数に影響を持つ要因であると考えられる。

もう一つの問題は、4つの雑誌からほぼ同数の論文を無作為抽出したことによりバイアスが生ずる可能性である。この問題は次の2つの論点に分けることができる。

(a) 被引用数分布における大きい歪度(skewness)を考えると、無作為抽出によりサンプルの被引用数分布が母集団のそれから外れないか?

(b) 規模(論文数)が異なる4つの雑誌からほぼ同数の論文を抽出することは妥当か?

(a)に関しては、無作為抽出によって分布が「系統的に」偏ることはない。しかし、歪度の大きい分布においては、非常に被引用数の多い少数の論文(“outlier”)がたまたま高い割合で抽出されるか、あるいは全く抽出されないことによって、サンプルの分布が「偶然的に」偏ることは起こりやすい。23の雑誌(1誌は全論文を抽出)でサンプルの被引用数(*C11*)平均値を母集団(雑誌の2000年の全原著論文)のそれと比較したところ、サンプルの平均値の方が高い雑誌が13誌、低い雑誌が10誌であり、高い方の1誌と低い方の2誌はその差が有意( $p < 0.05$ )であった。平均値は系統的に偏ってはいないが、有意差のある雑誌の比率(3/23)は5%より高く、これは、outlier論文の抽出が過剰あるいは過小であったためと考えられる。

(b)については、このようにサンプル論文を抽出した理由を1.4.1.3で述べたので、ここでは省略する。

## 4.5 本章の結論

6つの異なる分野において、論文の被引用数を説明する NBMR モデルを得た。4.4.1 及び 4.4.2 で述べたように、6つの分野におけるモデルは、選ばれた予測変数とその有意性の大きさにおいてある程度の共通性があった。被引用数の説明要因に関するこれまでのほとんどの研究は、1つの特定の分野の、または多くの分野を混合した論文集合に対しなされているので、被引用数の影響要因について異なる分野にわたって一般性を示した今回の結果は、この意味で独創性を持つと考えられる。

本章で示した NBMR 分析のもう一つの特徴は、目的変数に、比較的長期の被引用数を  $C6$  (論文発表後 6-7 年間) と  $C11$  (論文発表後 11-12 年間) の 2 通り設定したことである。2つの目的変数の結果がほぼ同様であったことは、得られた結果の安定性を示唆するものである。また、論文発表後 11-12 年間という期間は、これまでの類似の研究の中で最長である。表 4-8 に示した総合的重回帰研究のうち、Haslam et al. (2008) と Fu and Aliferis (2010) は 9-10 年、他は 6 年以下である。

NBMR の当てはまりは、pseudo  $R^2$  が 0.25~0.5 程度なので、非常に高くはなくまあまああのところである。今回用いた説明変数は全て、論文の質や内容に直接関係するものではない「外在的要因」であることを考えれば、これは予想される結果である。本研究の目的は、高い適合度のモデルを求めることではなく、これら外在的要因に基づく被引用数期待値のベースラインとなるモデルを探ることである。異なる分野にわたってある程度共通の予測要因が得られたことは、このベースライン設定が有望であることを示している。

今後の展開の一つは、このベースライン (外在的要因から予測される期待被引用数) からの個々の論文の実際の被引用数の外れの分析がある。この外れは、論文のどのような属性と関係しているのであろうか? それが論文の内在的要因 (質や内容を表す属性) と結びつけば、ここで提案したモデルを更に前進させることができるであろう。

もう一つの展開の方向は、被引用数と関係を持つ「内在的要因」(論文の質や内容とより密接に関係する指標) の探求である。これに関して Chen は、論文に含まれるアイデアの潜在的価値 (あるいは影響度) は、その論文がある研究分野の既存の知的構造を転換させる程度によって示されると考える構造変化 (structure variation) モデルを提唱した (Chen, 2012)。2.1.1.1 で述べたように、彼は、このモデルに基づき、ネットワーク理論を用いて構造変化の程度を測るいくつかの指標を定義し、それらの指標とその論文の被引用数との関係を論じている。

論文の被引用数にどんな要因が影響を及ぼすかについては、難しく複雑な問題が残っている。多くの研究がこの問題に取り組んでいるが、まだ明確な結論は得られていない。本章で述べた研究はこのテーマに一つの貢献をするものと考えている。

## 第5章

# 論文の引用持続性を示す指標の性質

### 【本章の要約】

長期間の引用累積曲線に基づいて論文の引用持続性を示す定量的な指標として Wang et al. (2015)により提案された Citation Delay ( $D$ ) の性質及び他の論文特性の間の関係を検討した。物性物理学、無機化学、電気・電子工学、生化学・分子生物学、生理学、消化器病学の6つの分野について、それぞれ4つの雑誌に2000年に発表された論文が15年間に得た引用データを用いて、以下の結果を得た。

- ・  $D$  はほぼ正規分布に従う。被引用数そのものに比べると分野間の格差は小さい。
- ・ 通算被引用数が多い論文ほど引用持続性が高い傾向がある。しかし、この関係は線形ではなく、 $D$ のある値で被引用数の平均値は最大となる。
- ・  $D$ を目的変数とし、多くの論文特性を説明変数とする重回帰分析の結果、Price 指数（参考文献中引用年齢が5年以内のもの比率）と  $D$ の間に明確な負の関係を見出した。また、図の数が多い論文は  $D$ が低く、表の数が多い論文は  $D$ が高い傾向がある。
- ・ それぞれの雑誌の中では高被引用論文の方が高い引用持続性を持つ傾向があるが、雑誌間で比べると、インパクトの高い雑誌には引用持続性の低い論文が含まれている割合が高い。

## 5.1 本章における研究目的

本章は、1.2において設定した3つの研究目標のうち、「研究目標3」について論述するものである。

研究目標3：論文の被引用数を観測する期間（引用計測期間）は、被引用数に影響を及ぼす外在的要因の一つであるが、この影響の検討は、研究目標2とは別途の方法によることが適当である。このため、論文の引用持続性を示す指標の分布の特徴、長期的被引用数との関係、外在的要因への依存性を、複数の分野において検討する。

第4章では、論文の被引用数がどのような外在的要因によって影響されるかについて論考し、いくつかの分野に共通して見られる傾向についての知見を得た。そこでは、同じサ

## 第5章 論文の引用持続性を示す指標の性質

サンプル中の論文が発表後 6-7 年間に得た被引用数(C6)と発表後 11-12 年間に得た被引用数(C11)を負の 2 項重回帰分析の目的変数としたが、それらに影響を与える要因(説明変数)に大きな差は見られなかった。

しかし、論文の被引用数の時間経過(引用の履歴)は、個々の論文によって大きく異なる。発表後 1~2 年間にその大部分が引用されてしまう論文、長期間コンスタントに引用される論文、数年間全く引用されずある時期になって突然引用され出す論文など様々である。従って、論文間の被引用数を比較するとき、どの程度の計測期間(引用ウィンドウ)をとるかにより結果は異なる。

引用計測期間による論文の被引用数への影響を明らかにするためには、論文による引用履歴の違いを客観的な(できれば定量的な)指標で表し、その指標と被引用数の関係を検討することが一つの方法であり、本章ではこの方法を採用する。

論文の引用履歴特性あるいは引用持続性に関する先行研究については 2.2 で詳しくレビューしたが、特に本章の研究に関係深い研究を以下に概括する。

論文をその引用持続性から遅延型(sleeping beauties)、早期型(flashes in the pan)、正規型(normal)などに分類する試みは従来からよくなされている(Garfield, 1980; Glänzel et al., 2003 ; van Raan, 2004 ; van Dalen and Henkens, 2005 ; Costas et al., 2010)。しかし、その定義や基準は人により異なるし、多様な引用履歴特性を少数のタイプに分類することにはやや無理がある。被引用半減期(被引用年齢分布の中央値)等いくつかの ageing パラメータは、被引用年齢分布全体のパターンを反映しているとは言えない。

遅延的に引用される論文の方が、早期に引用される論文より長期的に多数の引用を得る傾向があることが、いくつかの研究から示されている(Line, 1984; Aversa, 1985; Levitt and Thelwall, 2008, 2009a; Costas et al., 2010; Wang, 2013)。しかしこれらに使われた標本は、少数の高被引用論文に限定されていたり、いろいろな分野の論文が混在していたりして、どの分野にも共通の傾向であるかどうか明確でない。

長期的被引用数以外の論文特性と引用持続性との関係を比較した研究はごく少数である。van Dalen and Henkens (2005)と Costas et al. (2010)はいずれも、論文をその引用持続性によって 4 つのクラスに分け、それらのクラスに属する論文の特性を比較した。また、Wang et al. (2015)は、彼らが提案した引用持続性測度 Citation Delay に論文の学際性その他の特性が及ぼす影響を検討した。しかし、これらのいずれにおいても、研究の主要な目的は別にあり、論文の種々の属性と引用持続性の関係は比較的簡単にしか述べられていない。

このような事例はあるものの、論文の被引用数自体についての研究に比べると、引用持続性については系統的な知見の蓄積が不十分である。その本質的理由の一つは、論文の引用持続性を測る定量的な指標が確立されていないことにあると考えられる。上述の Citation Delay (Wang et al., 2015)は、被引用年齢分布の全域の情報をを用いた測度

## 第5章 論文の引用持続性を示す指標の性質

であり、引用持続性を示す測度としてこれまで提案されたものの中で最適と考えられるが、その性質についてはほとんど解明されていない。そこでこの章の研究では、この Citation Delay を論文の引用持続性指標として用い、この指標の性質を次の3つの観点から解明する。

- (1) Citation Delay の分布の特徴を検討する。
- (2) 異なる分野において、この持続性指標と論文の被引用数の関係を調査し、早期に引用される論文と遅延的に引用される論文の間で、長期的被引用数に傾向の違いがあるか否かを検討する。
- (3) 異なる分野において、論文の諸特性と引用持続性の関係を調べ、分野を超えた共通の傾向が見られるか否かを検討する。

### 5.2. データと方法

#### 5.2.1 用いる持続性指標

本論文で引用持続性の指標として用いる Wang et al (2015)による Citation Delay (以下これを記号  $D$  で表す) については 2.2.4 で簡単に述べたが、ここで改めて定義する。

ある論文の発表された年を  $j = 0$ 、引用を観測した最後の年を  $T$  としたとき、 $j = 0, 1, 2, \dots, T$  の各年にこの論文が得た引用数を  $c(j)$  とする。すると、発表年から  $j = t$  までの累積被引用数  $C(t)$  は次式で与えられる。

$$C(t) = \sum_{j=0}^t c(j) \quad [5-1]$$

$j = t$  までに得た累積被引用比率  $x(t)$  は

$$x(t) = \frac{C(t)}{C(T)} \quad [5-2]$$

である。 $j = 0$  から  $j = T$  までの全被引用数  $C(T)$  を、以下では通算被引用数と呼ぶ。 $j = T$  におけるこの論文の  $D$  を次式で定義する。

$$D = 1 - \frac{1}{T} \sum_{t=0}^{T-1} x(t) \quad [5-3]$$

図 5-1 は、横軸に  $t (0 \leq t \leq T-1)$ 、縦軸に  $x(t)$  をとって、 $D$  の意味を模式的に説明したものである。この長方形の面積を 1 としたとき、影を付けた部分の面積が  $D$  である。

論文の引用すべてが発表年 ( $t = 0$ ) になされ、その後全く引用がなかったとすれば、図 5-1 の  $t = 0$  で  $x(t)$  は 1 に達し、以後そのままなので、影部分の面積(従って  $D$ ) は

0である。逆に、 $t=0$  から  $t=T-1$  まで全く引用がなく、 $t=T$ である数の引用がなされた場合、図5-1の面積のすべてを影が占め、 $D=1$ である。それ故、 $D$ は  $0 \leq D \leq 1$  の範囲にあり、早期の引用が多いほど影の部分の面積（従って  $D$ ）は小さく、遅延引用が多いほど影の部分の面積（従って  $D$ ）は大きくなる。すなわち、 $D$ は被引用年齢分布の全域の情報を用いた持続性の指標といえることができる。

$D$ は図5-1のグラフの相対的形状にのみ依存し、通算被引用数  $\alpha(T)$ には依存しない。但し、 $\alpha(T)=0$ の場合  $D$ は定義できず、 $\alpha(T)$ が非常に小さい場合もあまり意味がない。ここでは、主として15年間の  $\alpha(T)$ が5以上の論文に対して分析を行う。

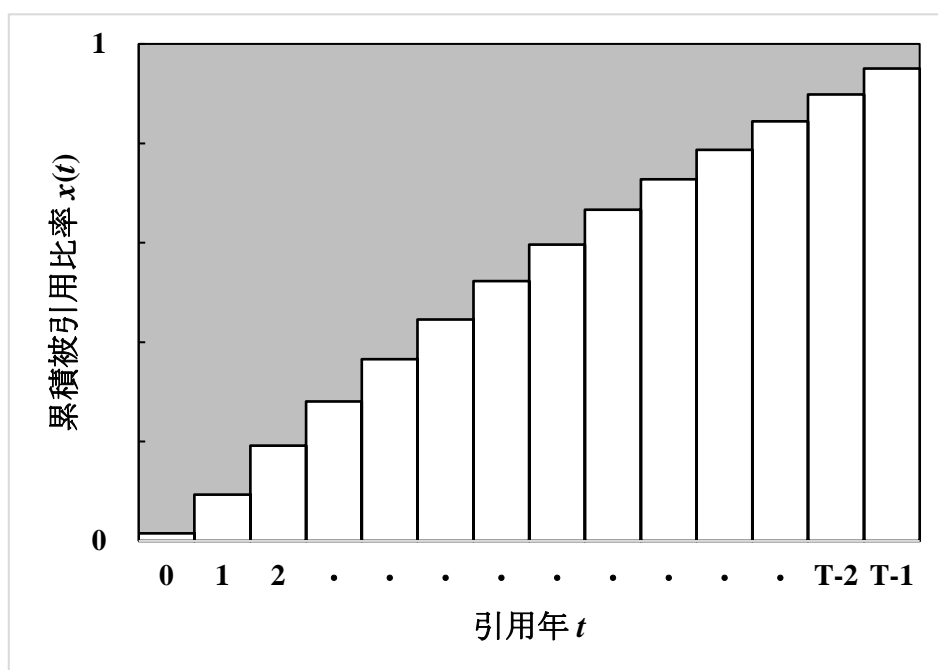


図5-1 Citation Delay ( $D$ )の意味を示す模式図

長方形の面積を1とするとき、影の部分の面積が  $D$ になる。

### 5.2.2 用いるデータ

ターゲットとする分野として、Web of Science（以下 WoS と略記）の主題カテゴリー（Subject Categories）から次の6つを選んだ。以降では、これらの分野を下記の下線部で略記する。

- ・物理学：物性物理学(condensed-matter physics)
- ・化学：無機・核化学(inorganic chemistry)
- ・工学：電気・電子工学(electric and electronic engineering)
- ・生物学：生化学・分子生物学(biochemistry and molecular biology)



## 第5章 論文の引用持続性を示す指標の性質

- ・基礎医学：生理学(physiology)
- ・臨床医学：消化器病学(gastroenterology)

これらの分野のそれぞれから4誌ずつ、表5-1に示す雑誌を選んだ。分野、雑誌をこのように選択した理由については、1.4.1に詳しく述べた。

これらの24誌から、2000年に発表された正規論文（WoSのデータタイプが“article”であるもの）を分析の対象とした。但し、“article”と同時に“proceedings paper”にも分類されているもの、及び著者が存在しないものを除外した。表5-1には各雑誌から採られた論文数も示す。

2015年3月31日に、WoSからこれらの論文の引用データをダウンロードし、2000年( $t=0$ )から2014年( $t=T=14$ )までの各年の被引用数を記録した。

本章での分析はそれぞれの分野単位で行うほか、表5-1の「雑誌コード」欄にコードを示した12の雑誌（各分野2誌ずつ）については、雑誌単位でも行った。

### 5.2.3 $D$ の算出

24誌に含まれる18,702論文中、331論文は2000–2014年の期間に1回も引用されておらず、通算被引用数 $\alpha(T)$ が0である。これらを除いた18,371論文に対して、5.2.1で導入した $D$ を、式[5-1]～式[5-3]に従い計算した。但し、このうち1,556論文は $\alpha(T)$ が1～4で、引用持続性の測定があまり意味を持たないと考えられるので、以下では $\alpha(T) \geq 5$ の16,815論文を主な分析の対象とする。

$D$ がどのような分布に従うかについてはほとんど知られていない（Wang et al., (2015)に“citation delay is roughly normally distributed”という記述があるが、それを裏付けるデータは示されていない）。本章では、6つのそれぞれの分野、及び表5-1の「雑誌コード」欄にコードを示した12の雑誌の中で $D$ の分布を調べ、それを正規分布と比較した。

### 5.2.4 引用持続性と通算被引用数の間の関係

6つのそれぞれの分野及び12のそれぞれの雑誌の中で、 $\alpha(T) \geq 5$ の論文群に対して、 $D$ と $\alpha(T)$ の間の相関を調べた。 $\alpha(T)$ の分布は極めて歪度が大きいので、これを対数変換した $\log(\alpha(T)+1)$ を用いた。Pearson積率相関を求めるほか、非線形な関係も分析した。。

表 5-1 選定した主題分野、雑誌、及び抽出論文数

主題分野	雑誌名 <sup>a</sup>	発行国	論文数	雑誌コード <sup>b</sup>
物性物理学 (物性)	European Physical Journal B	DEU	538	
	Journal of Physics - Condensed Matter	GBR	813	CondMat-1
	Physica B	NLD	148	
	Physical Review B	USA	4738	CondMat-2
無機・核化学 (無機)	Inorganic Chemistry	USA	931	Inorg-1
	Inorganica Chimica Acta	CHE	546	
	Journal of the Chemical Society - Dalton Transactions	GBR	682	Inorg-2
	Transition Metal Chemistry	NLD	139	
電気・電子工学 (電気)	IEE Proceedings - Circuits, Devices and Systems	GBR	52	
	IEEE Transactions on Circuits and Systems I	USA	218	
	IEEE Transactions on Microwave Theory and Techniques	USA	295	Elec-1
	Signal Processing	NLD	178	Elec-2
生化学・分子生物学 (生化)	European Journal of Biochemistry	GBR	788	Biochem-1
	Journal of Biochemistry (Tokyo)	JPN	275	
	Journal of Biological Chemistry	USA	5504	Biochem-2
	Journal of Molecular Biology	USA	875	
生理学 (生理)	Journal of General Physiology	USA	110	Physiol-1
	Japanese Journal of Physiology	JPN	72	
	Journal of Physiology - London	GBR	472	Physiol-2
	Pflugers Archive European Journal of Physiology	DEU	238	
消化器病学 (消化器)	American Journal of Gastroenterology	USA	430	Gastro-1
	Gastroenterology	USA	259	
	Gut	GBR	277	Gastro-2
	Journal of Gastroenterology	JPN	124	

a) 2000年時点の雑誌名（その後変更されたものあり）。

b) コードが付与された12誌は雑誌レベルの分析に使用(5.2.2を参照)。以降の本文ではこれらのコードを使用。

### 5.2.5 論文の諸特性から $D$ を説明する重回帰分析

どのような論文特性が  $D$  と強い関係を持つかを検討するために、6つの分野別に、 $C(T) \geq 5$  の論文を対象に、 $D$  を目的変数、表 5-2 に示す諸特性を説明変数とする重回帰分析を行った。重回帰分析に先立つ予備分析として、 $D$  と各説明変数の間の Pearson 相関係数を求めた。

これらの説明変数には、第4章で述べた被引用数そのものに影響を及ぼす要因に関する研究で用いた外在的要因の他に、通算被引用数  $C(T)$  と論文が発表された雑誌のインパクトファクター ( $JIF$ ) を加えた。これは、 $D$  とこれらの変数との間に相関があるという結果が得られたためである (5.3.3 及び 5.3.4 を参照)。表 5-2 の第4列

## 第5章 論文の引用持続性を示す指標の性質

に各説明変数の値を得る方法を示すが、より詳しくは 4.2.3 を参照されたい。

表 5-2  $D$  を予測する重回帰分析に用いる説明変数

説明変数	定義	対数変換	データの所得方法
<i>Authors</i>	論文の著者数	なし	WoSのAUフィールドのエントリーをカウント
<i>Insts</i>	著者の所属する機関(著者所属機関)の数	なし	WoSのC1フィールドのエントリーをカウント
<i>Countries</i>	著者所属機関が属する国(著者所属国)の数	なし	WoSのC1フィールド中の異なる国名をカウント
<i>Refs</i>	論文に引用された参考文献の数	なし	WoSのNRフィールドの値
<i>Price</i>	Price指数(参考文献のうち、その発表年が論文の発表年の5年前以内のものパーセンテージ)	なし	WoSのCRフィールドのデータから計算
<i>Length</i>	ページあたり文字数を規格化した論文のページ数	なし	原論文から数えたページ数をページあたり文字数で規格化(6,400文字/ページとする)
<i>Figures</i>	論文中の図の数	なし	原論文からカウント (figure, chart, diagram, picture等を含む)
<i>Tables</i>	論文中の表の数	なし	原論文からカウント
<i>Eqs</i>	論文中の番号付き数式の数	$\log(Eqs + 1)$	原論文からカウント
<i>Age</i>	第一著者の活動期間(最初の論文の発表年から2000年までの経過年)	$\log(Age + 1)$	WoSの著者名サーチ(発表年期間1970-2000)の結果を分析 <sup>a</sup>
<i>RatePubl</i>	第一著者の活動機関内の年平均発表論文数	$\log(RatePubl + 0.1)$	WoSの著者名サーチ(発表年期間1970-2000)の結果を分析 <sup>a</sup>
<i>MedCite</i>	活動期間中の発表論文の年あたり被引用数(2000年までの中央値)	$\log(MedCite + 0.01)$	上記の著者名サーチから得られた論文の引用データに基づき計算
$C(T)$	論文が2014年末までに得た引用数	$\log(C(T) + 1)$	WoSから入手
<i>JIF</i>	論文が発表された雑誌のインパクトファクター(2001年と2002年の平均)	なし	Journal Citation Reportsから入手

a) 詳細については 3.3 を参照。

説明変数の分布を調べた結果から、過半の分野で歪度が 2 以上あるいは平均値と中央値の比が 1.5 以上のいずれかに該当する  $C(T)$ 、*Eqs*、*Age*、*RatePubl*、*MedCite* については対数変換を行った。歪度は Excel の SKEW 関数により計算したが、これは母集団統計量に対する不偏推定量の式、 $\frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s}\right)^3$  に従う。いずれの変数についても値が 0 の論文があるため、 $C(T)$ 、*Eqs*、*Age* については 1、*RatePubl* については 0.1、*MedCite* については 0.01 を加えた値の対数をとった(表 5-2 の第 3 列を参照)。これらの加算値は、それぞれの変数の分布を考慮して定めた。

重回帰分析には、24 の雑誌の全論文(articles)ではなく、各誌からそれぞれ 60 以下の論文を無作為抽出したものをサンプルとして用いた。このサンプル論文は、第 4 章の重回帰分析に用いたもの(4.2.1 及び表 1-1 を参照)とほぼ同じであるが、 $C(T) \geq 5$  の論文に限定しているのもそれよりやや少なくなる。サンプルとして用いた論文数を表 5-3 に示す。

## 第5章 論文の引用持続性を示す指標の性質

重回帰分析には SPSS/PASW statistics Base 17.0 を用い、ステップワイズ法により有意な説明変数を選択した。

表 5-3 重回帰分析のサンプルに用いた論文数

分野	雑誌	サンプル論文数
物性	European Physical Journal B	49
	Journal of Physics - Condensed Matter	42
	Physica B	39
	Physical Review B	47
無機	Inorganic Chemistry	50
	Inorganica Chimica Acta	53
	Journal of the Chemical Society - Dalton Transactions	49
	Transition Metal Chemistry	40
電気	IEE Proceedings - Circuits, Devices and Systems	25
	IEEE Transactions on Circuits and Systems I	50
	IEEE Transactions on Microwave Theory and Techniques	37
	Signal Processing	33
生化	European Journal of Biochemistry	57
	Journal of Biochemistry (Tokyo)	54
	Journal of Biological Chemistry	60
	Journal of Molecular Biology	60
生理	Journal of General Physiology	34
	Japanese Journal of Physiology	60
	Journal of Physiology - London	55
	Pflugers Archive European Journal of Physiology	50
消化器	American Journal of Gastroenterology	47
	Gastroenterology	56
	Gut	52
	Journal of Gastroenterology	44

## 5.3 結果

### 5.3.1 通算被引用数 $\alpha(T)$ の分布

本論である  $D$  の性質を論ずる前に、通算被引用数  $\alpha(T)$  の分布の概要を記す。

6つの分野と各分野から2誌ずつ選んだ12の雑誌について、 $\alpha(T)$ の主要な統計量とともに、 $\alpha(T) \geq 1$  及び  $\alpha(T) \geq 5$  の論文数を表 5-4 に示す。よく知られているように、論文の被引用数は、同じ分野、同じ雑誌の中でも極めて歪度の大きい分布を

## 第5章 論文の引用持続性を示す指標の性質

示すが、ここに示す  $C(T)$  の分布もその例に漏れない。

表 5-4  $C(T)$  に関する統計

(a) 分野レベル

分野	論文数	平均	標準偏差	中央値	歪度	$C(T) \geq 1$ の論文数	$C(T) \geq 5$ の論文数
物性	6,237	30.0	72.8	15	33.1	6,028	5,097
無機	2,298	28.9	33.5	19	4.8	2,278	2,061
電気	743	19.6	36.5	9	5.9	698	517
生化	7,442	63.5	78.1	44	12.5	7,432	7,327
生理	892	38.9	51.9	26	7.4	876	808
消化器	1,090	60.9	81.9	38	4.0	1,059	1,005

(b) 雑誌レベル

雑誌	論文数	平均	標準偏差	中央値	歪度	$C(T) \geq 1$ の論文数	$C(T) \geq 5$ の論文数
CondMat-1	813	15.5	22.5	9	4.0	768	570
CondMat-2	4,738	34.2	81.9	18	30.4	4,610	4,046
Inorg-1	931	36.5	35.1	26	2.5	929	882
Inorg-2	682	30.2	37.8	20	6.8	680	626
Elec-1	295	21.9	35.0	12	4.6	284	228
Elec-2	178	9.9	15.0	5	3.3	161	97
Biochem-1	788	32.4	45.5	23	12.5	785	754
Biochem-2	5,504	70.5	67.4	51	3.2	5,501	5,469
Physiol-1	110	50.9	37.2	44.5	1.1	109	107
Physiol-2	472	50.3	64.1	34	6.9	470	459
Gastro-1	430	43.3	46.6	29.5	2.9	408	385
Gastro-2	277	65.4	78.0	44	5.1	275	269

そこで、対数変換した  $\log[C(T)+1]$  を考え ( $C(T)=0$  の論文があるので、これに 1 を加えたものの対数を用いる)、その分布統計量を表 5-5 に示す。 $C(T)$  では平均値と中央値が大きく異なる (平均値は中央値の 1.5~2 倍) が  $\log[C(T)+1]$  では両者はほとんど一致する。また、 $C(T)$  に比べて  $\log[C(T)+1]$  の歪度は著しく減少している。これらのことから、 $\log[C(T)+1]$  は平均値の周りにほぼ対称に分布していると推測できる。

$\log[C(T)+1]$  の分布が正規分布に近いかどうかを見るための Q-Q plot を図 5-2 に示す。Normal theoretical quantile が -2 から +2 の範囲 (ほぼ 2.5~97.5% 区間に相当) では、消化器分野がやや湾曲している以外はほぼ直線状である。従って、以下の分析では、正規性を仮定した統計解析 (Pearson 積率相関、線形重回帰分析) を  $\log[C(T)+1]$  に適用することとする。

表 5-5  $\log[C(T)+1]$  に関する統計

(a) 分野レベル

分野	論文数	平均	標準偏差	中央値	歪度
CondMat	6,237	1.20	0.51	1.2	-0.1
Inorg	2,298	1.29	0.42	1.3	-0.3
Elec	743	1.00	0.51	1.0	0.1
Biochem	7,442	1.64	0.39	1.7	-0.3
Physiol	892	1.38	0.46	1.4	-0.5
Gastro	1,090	1.52	0.54	1.6	-0.6

(b) 雑誌レベル

雑誌	論文数	平均	標準偏差	中央値	歪度
CondMat-1	813	0.98	0.46	1	0.0
CondMat-2	4,738	1.26	0.50	1	-0.2
Inorg-1	931	1.42	0.39	1	-0.3
Inorg-2	682	1.32	0.39	1	-0.2
Elec-1	295	1.09	0.48	1	0.0
Elec-2	178	0.79	0.46	1	0.2
Biochem-1	788	1.37	0.36	1	-0.2
Biochem-2	5,504	1.71	0.36	2	-0.2
Physiol-1	110	1.58	0.38	1.7	-1.0
Physiol-2	472	1.54	0.39	2	-0.3
Gastro-1	430	1.41	0.53	1.5	-0.9
Gastro-2	277	1.63	0.44	2	-0.6

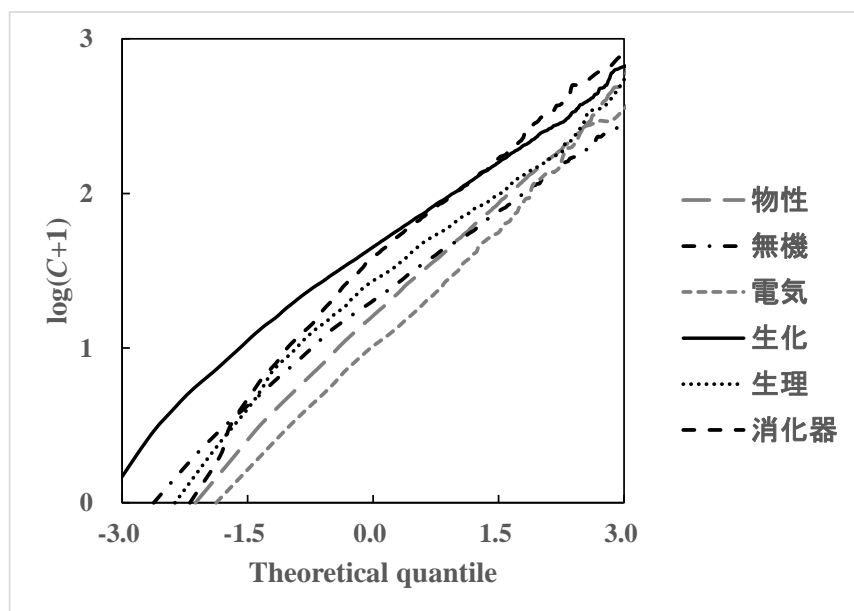


図 5-2  $\log[C(T)+1]$  の分布に対する Q-Q プロット ( $\alpha(T) \geq 1$  の論文を対象)

### 5.3.2 $D$ の分布

#### 5.3.2.1 分布の概要

$\alpha(T) \geq 5$  の論文群に対し、6つの分野について式[5-3]により計算した  $D$  の分布の概要を図 5-3 に示す。 $D$  の値が 0.4~0.5 付近を中心としてほぼ左右対称な分布になっている。より詳細に、6つの分野及び各分野からそれぞれ2誌ずつ選んだ12誌のそれぞれについて、 $D$  の以下の統計量を表 5-6 ( $\alpha(T) \geq 1$  の論文群) と表 5-7 ( $\alpha(T) \geq 5$  の論文群) に示した。

- ・ 平均値(Mean)
- ・ 標準偏差(SD)
- ・ 第1、第2、第3四分位数(P25, P50, P75)
- ・ 四分位数間の差(P50-P25, P75-P50)

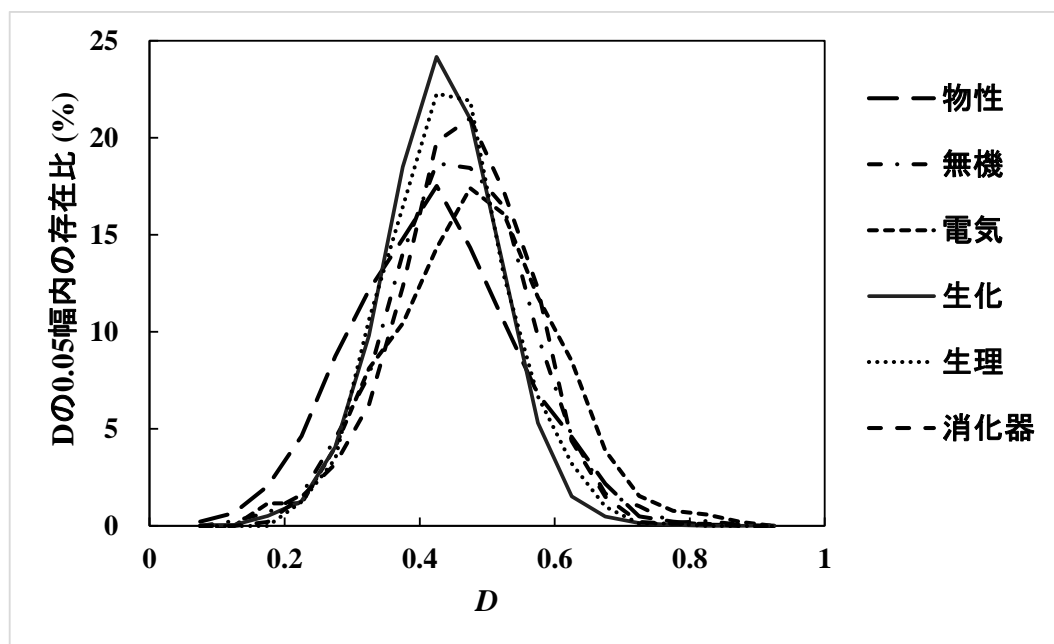


図 5-3  $D$  の分布 ( $\alpha(T) \geq 5$  の論文を対象)

第5章 論文の引用持続性を示す指標の性質

表 5-6  $D$  に対する統計量 ( $C(T) \geq 1$  の論文を対象)

(a) 分野レベル

分野	$n$	平均	標準偏差	P25	P50	P75	P50-P25	P75-P50
物性	6,028	0.411	0.139	0.321	0.411	0.499	0.090	0.088
無機	2,278	0.450	0.119	0.380	0.452	0.523	0.073	0.070
電気	698	0.462	0.157	0.359	0.473	0.555	0.114	0.081
生化	7,432	0.431	0.087	0.376	0.434	0.488	0.058	0.054
生理	876	0.441	0.102	0.382	0.443	0.500	0.061	0.057
消化器	1,059	0.462	0.106	0.400	0.465	0.531	0.065	0.066

(b) 雑誌レベル

雑誌	$n$	平均	標準偏差	P25	P50	P75	P50-P25	P75-P50
CondMat-1	768	0.422	0.157	0.324	0.429	0.525	0.104	0.097
CondMat-2	4,610	0.408	0.131	0.321	0.409	0.490	0.088	0.081
Inorg-1	929	0.453	0.106	0.390	0.455	0.524	0.065	0.069
Inorg-2	680	0.434	0.112	0.367	0.435	0.503	0.068	0.068
Elec-1	284	0.449	0.143	0.361	0.462	0.536	0.101	0.074
Elec-2	161	0.459	0.191	0.343	0.464	0.585	0.121	0.121
Biochem-1	785	0.443	0.105	0.375	0.450	0.510	0.075	0.061
Biochem-2	5,501	0.430	0.083	0.377	0.433	0.484	0.056	0.051
Physiol-1	109	0.432	0.075	0.389	0.423	0.478	0.034	0.054
Physiol-2	470	0.444	0.093	0.387	0.446	0.499	0.059	0.053
Gastro-1	408	0.468	0.112	0.399	0.476	0.545	0.077	0.069
Gastro-2	275	0.463	0.102	0.404	0.472	0.533	0.068	0.061

表 5-7  $D$  に対する統計量 ( $C(T) \geq 5$  の論文を対象)

(a) 分野レベル

分野	$n$	平均	標準偏差	P25	P50	P75	P50-P25	P75-P50
物性	5,097	0.420	0.119	0.338	0.419	0.498	0.081	0.079
無機	2,061	0.456	0.105	0.388	0.457	0.522	0.069	0.065
電気	517	0.480	0.119	0.399	0.484	0.554	0.085	0.069
生化	7,327	0.432	0.084	0.378	0.435	0.488	0.057	0.054
生理	808	0.443	0.088	0.385	0.443	0.496	0.058	0.053
消化器	1,005	0.464	0.095	0.402	0.466	0.528	0.064	0.062

(b) 雑誌レベル

雑誌	$n$	平均	標準偏差	P25	P50	P75	P50-P25	P75-P50
CondMat-1	570	0.437	0.130	0.352	0.436	0.531	0.084	0.095
CondMat-2	4,046	0.417	0.115	0.337	0.417	0.491	0.080	0.074
Inorg-1	882	0.458	0.098	0.397	0.458	0.526	0.062	0.068
Inorg-2	626	0.438	0.099	0.374	0.437	0.504	0.064	0.066
Elec-1	228	0.460	0.110	0.380	0.469	0.535	0.089	0.066
Elec-2	97	0.495	0.138	0.411	0.490	0.588	0.079	0.099
Biochem-1	754	0.448	0.100	0.383	0.454	0.513	0.071	0.059
Biochem-2	5,469	0.430	0.081	0.377	0.433	0.484	0.055	0.051
Physiol-1	107	0.430	0.072	0.389	0.423	0.475	0.034	0.052
Physiol-2	459	0.448	0.088	0.390	0.448	0.500	0.057	0.052
Gastro-1	385	0.471	0.100	0.403	0.478	0.544	0.075	0.066
Gastro-2	269	0.468	0.094	0.405	0.472	0.535	0.068	0.063



これらの表から次のことが判る。

- (1) 平均値、中央値については、物性、無機、電気の3分野では  $\alpha(T) \geq 5$  の論文群の方が  $\alpha(T) \geq 1$  の論文群よりやや大きい。が、生化、生理、消化器の3分野では  $\alpha(T) \geq 5$  と  $\alpha(T) \geq 1$  の差はほとんどない。
- (2) 標準偏差は 0.08~0.15 程度の大きさであるが、 $\alpha(T) \geq 5$  の論文群では  $\alpha(T) \geq 1$  の論文群に比べて明らかに減少している（減少率は、分野、雑誌により 2~28%）。同様に、四分位数間の差（P50-P25 及び P75-P50）も、ほとんどの分野、雑誌で  $\alpha(T) \geq 5$  の論文群の方が小さい。つまり、通算被引用数がある程度高い論文に限定すると、 $D$  の分布の幅が狭まる。
- (3) どの分野、雑誌でも、 $D$  の平均値と中央値はかなり近い（特に  $\alpha(T) \geq 5$  の論文群では）。また、P25 と P50 の間隔と P50 と P75 の間隔の間に大きな差がない。これらのことから、分布が平均値の周りにほぼ対称であるという図 5-3 からの印象が裏付けられる。

### 5.3.2.2 分布の正規性の検証

表 5-6、表 5-7 から、P50-P25、P75-P50 の値と標準偏差(SD)の比をとってみると 0.46~0.73 の範囲にあり、正規分布におけるこの比の理論値 0.674 に近い。 $\alpha(T) \geq 5$  の論文群について、正規分布との比較を示す Q-Q plot を図 5-4 に示す。プロットはどの分野も直線に近い。 $\alpha(T) \geq 1$  の論文群に対する Q-Q plot（図は示していない）も概ね直線上だが、 $D$  が 0.6 より大きい領域で勾配がやや急になる。

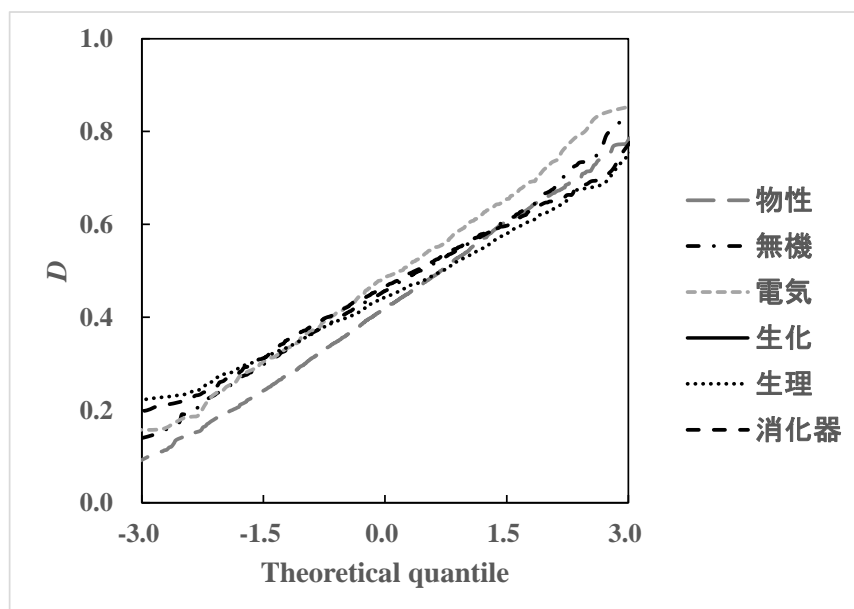


図 5-4  $D$  の分布に対する Q-Q プロット ( $\alpha(T) \geq 5$  の論文を対象)

## 第5章 論文の引用持続性を示す指標の性質

以上のことから、 $D$ の分布はかなり正規分布に近く、特に、ある程度以上の被引用数を持つ論文では、この特徴は明確である。 $D$ のこの性質は統計学的指標として好ましいと言える。

### 5.3.2.3 分野間の平均値の差

論文の被引用数分布は分野によって大きな開きがあることが知られており、表 5-4、表 5-5 にもそのことが現れているが、 $D$ の分布についてはどうであろうか？ 図 5-5a は  $D$ 、図 5-5b は  $\log[C(D)+1]$ の分布を箱ひげ図によって示したものである（どちらも  $C \geq 1$  の論文を対象）が、分野間の違いがより視覚的に比較できる。 $\log[C(D)+1]$ に比べ  $D$ の分布は、分野間でかなり重なり合っているように見える。

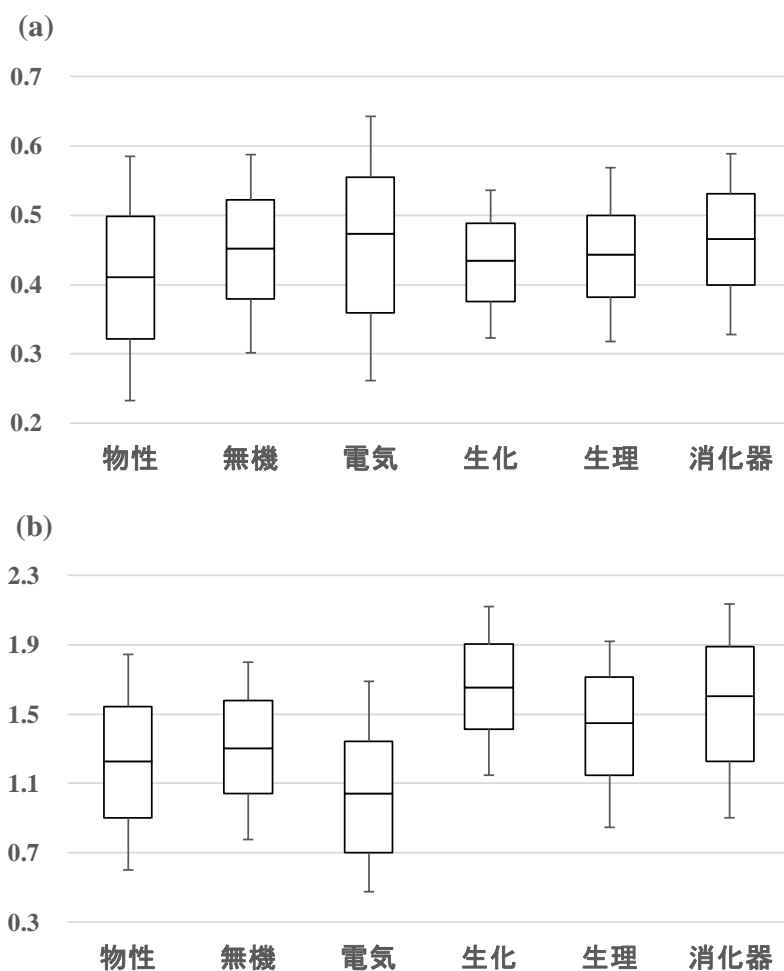


図 5-5 分野による分布範囲の違い (a)  $D$ の分布 (b)  $\log[C(D)+1]$ の分布

ボックス部の中央線は中央値(第2四分位値)、両端は第1及び第3四分位値。  
ボックス外側のバーの両端は10パーセンタイル値と90パーセンタイル値。

## 第5章 論文の引用持続性を示す指標の性質

このことを確かめるために、 $D$  と  $\log[C(T)+1]$  のそれぞれに対し、すべての分野対の間の分布の重なりを次の方法により計算した：

- (1) ある分野 (A とする) における P25 と P75 を求める ( $D$  と  $\log[C(T)+1]$  のそれぞれについて)
- (2) A 以外の各分野 (B とする) に対し、A の [P25, P75] 区間と重なり合うパーセンタイル区間を計算する。

たとえば、A を物性、B を無機としたとき、A における  $D$  の [P25, P75] は [0.321, 0.499] であり、B における  $D$  の分布で 0.321 と 0.499 に当たるのはそれぞれ 12.7 パーセンタイル、66.1 パーセンタイルである。従って、B の A に対する重なりは 53.4 パーセンタイルになる。

このようにして、30 の分野対 (6 つの A に対しそれぞれ 5 つの B がある) に対する重なりパーセンタイル区間が得られる。図 5-6 に示す  $D$  と  $\log[C(T)+1]$  に対する重なり値の累積分布から、 $D$  が  $\log[C(T)+1]$  に比べて明らかに分野間の重なりが大きいことが判る。

以上のことから、引用持続性に関しては、被引用数そのものほど明瞭な分野間格差は見られないと言える。また、一般に論文の被引用数は生物・医学系、物理・化学系、工学系の順に高い (今回の  $C(T)$  でも同様) が、 $D$  ではそのような傾向はない。

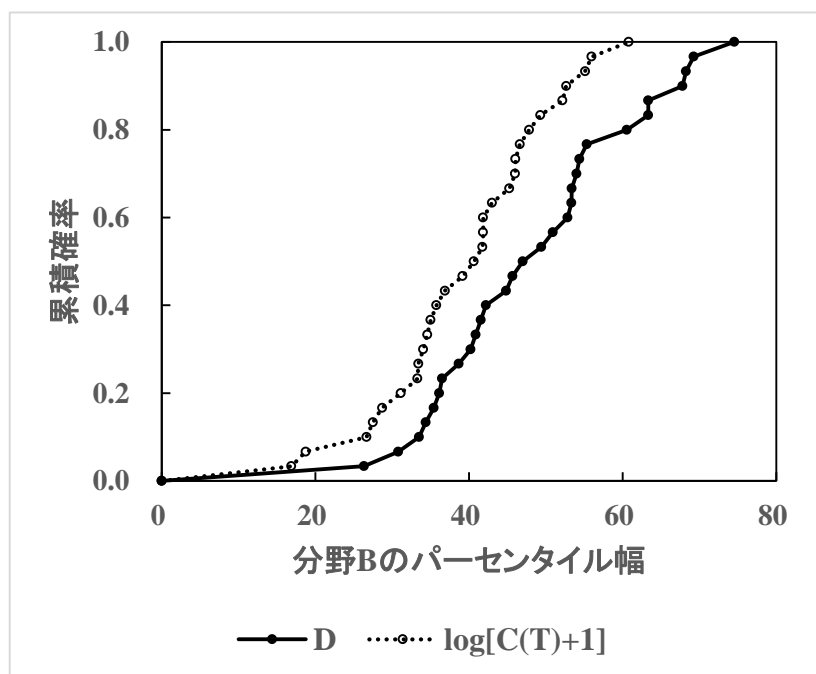


図 5-6 分野 A の [P25, P75] 範囲に含まれる分野 B のパーセンタイル幅の分布

### 5.3.3 $D$ と $\log[C(T)+1]$ の間の関係

$C(T) \geq 5$  の論文について算出した  $\log[C(T)+1]$  と  $D$  の間の Pearson 積率相関係数  $r(D-C)$  を表 5-8 に示す。 $r(D-C)$  はすべて 0.15~0.5 の範囲にあり、95%信頼区間に 0 を含むものはない。Spearman 順位相関係数も、ほとんど同じ値であった。つまり、通算被引用数が多い論文ほど引用持続性が高い傾向があるように見える。

表 5-8  $D$  と  $\log[C(T)+1]$  の間の Pearson 相関係数  $r(D-C)$ .

(a) 分野レベル

分野	$n$	$r(D-C)$	95% CI
物性	5,097	0.325	[0.300, 0.350]
無機	2,061	0.167	[0.124, 0.208]
電気	517	0.259	[0.177, 0.338]
生化	7,327	0.288	[0.267, 0.309]
生理	808	0.291	[0.227, 0.353]
消化器	1,005	0.198	[0.138, 0.256]

(b) 雑誌レベル

雑誌	$n$	$r(D-C)$	95% CI
CondMat-1	570	0.330	[0.254, 0.401]
CondMat-2	4,046	0.366	[0.339, 0.392]
Inorg-1	882	0.271	[0.209, 0.331]
Inorg-2	626	0.159	[0.082, 0.235]
Elec-1	228	0.260	[0.135, 0.377]
Elec-2	97	0.321	[0.130, 0.489]
Biochem-1	754	0.323	[0.258, 0.386]
Biochem-2	5,469	0.324	[0.300, 0.348]
Physiol-1	107	0.471	[0.309, 0.607]
Physiol-2	459	0.346	[0.263, 0.424]
Gastro-1	385	0.293	[0.199, 0.382]
Gastro-2	269	0.210	[0.092, 0.321]

しかし、もう少し詳しく見ると、この関係は線形ではない。表 5-9 は、 $D$  を 0.05 間隔に刻んだ区間ごとの各分野における  $\log[C(T)+1]$  の平均値である。 $D$  のある値で  $\log[C(T)+1]$  の平均値は最大となることが読みとれる。物性分野と生化分野について、これらの関係を散布図で示したものが図 5-7 である。平均値を示す位置（図中に●で示す）から、 $\log[C(T)+1]$  が  $D$  に対して非線形に変化することが解る。つまり、ごく弱い傾向ではあるが、よく引用される論文はあまり早期型でも遅延型でもなく、どの時期にも均等に引用される。

また、図 5-7 を見ると、低被引用部分（グラフの下部）では  $D$  が全域に拡がっているが、高被引用側（グラフの上部）に移るにつれて  $D$  の存在域が狭くなっている

第5章 論文の引用持続性を示す指標の性質

のが解る。5.3.2.1で  $C(T) \geq 1$  の論文群と  $C(T) \geq 5$  の論文群の比較から、「通算被引用数がある程度高い論文に限定すると、 $D$ の分布の幅が狭まる」と述べたが、これと符合する結果である。

表 5-9  $D$  と  $\log[C(T)+1]$  の間の非線形な関係

$D$ の範囲	$\log[C(T)+1]$ の平均 <sup>a</sup>					
	物性	無機	電気	生化	生理	消化器
$0.05 \leq D < 0.1$	0.88			0.78		
$0.1 \leq D < 0.15$	0.96	0.92		0.95		
$0.15 \leq D < 0.2$	1.06	0.99	0.97	1.11		1.02
$0.2 \leq D < 0.25$	1.08	1.07	0.97	1.21	1.08	1.17
$0.25 \leq D < 0.3$	1.18	1.14	1.00	1.34	1.24	1.26
$0.3 \leq D < 0.35$	1.25	1.28	1.07	1.49	1.31	1.44
$0.35 \leq D < 0.4$	1.34	1.35	1.15	1.63	1.42	1.57
$0.4 \leq D < 0.45$	1.42	1.42	1.19	1.69	1.48	1.62
$0.45 \leq D < 0.5$	1.47	1.45	1.39	1.75	1.57	1.69
$0.5 \leq D < 0.55$	1.49	1.42	1.29	1.75	1.60	1.66
$0.55 \leq D < 0.6$	1.53	1.46	1.37	1.77	1.63	1.71
$0.6 \leq D < 0.65$	1.48	1.39	1.36	1.71	1.59	1.70
$0.65 \leq D < 0.7$	1.50	1.31	1.41	1.56	1.10	1.50
$0.7 \leq D < 0.75$	1.63	1.26	1.43	1.55		1.59
$0.75 \leq D < 0.8$	1.13	0.91	0.93	0.85		
$0.8 \leq D < 0.85$	0.86	0.88	1.00			
$0.85 \leq D < 0.9$	1.18					

a) 2論文以上が含まれる場合のみ平均値を記載。

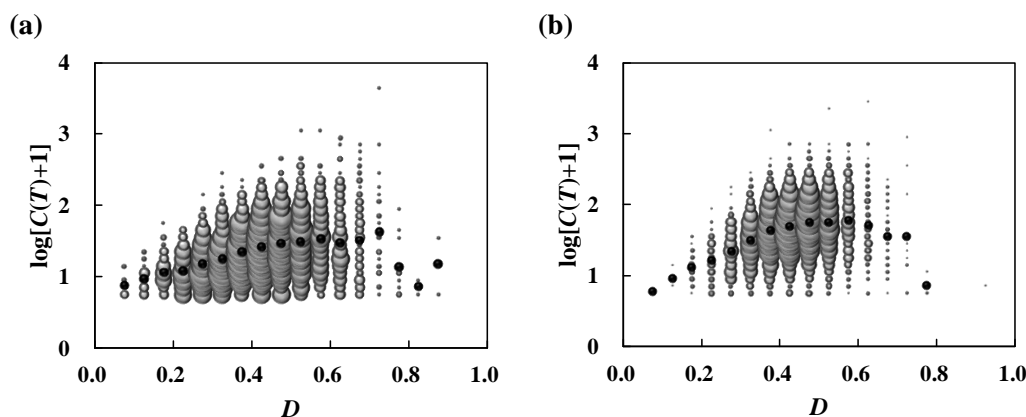


図 5-7  $D$  の変化に対する  $\log[C(T)+1]$  の変化の非線形性 (a) 物性分野 (b) 生化分野

バブルの大きさは、各細分領域に存在する論文の数に対応。

ドット(●)はそれぞれの  $D$  区間における  $\log[C(T)+1]$  の平均。

## 第5章 論文の引用持続性を示す指標の性質

なお、 $\log[C(T)+1]$ の値により区切った区間ごとの  $D$  の平均値の変化は山型にはならず、単調に増加する。

### 5.3.4 雑誌による引用持続性の違い

前節で被引用数の多い論文は引用持続性も高い傾向があることを示したが、それでは、インパクトの高い雑誌の論文は平均的に持続性が高いだろうか？ 表 5-10 に、各分野の各 4 誌に対する  $D$  の平均値 ( $C \geq 1$  の全論文に対する)、及びそれらの雑誌のインパクトファクター( $JIF$ )を示す。2000 年発表の論文を対象としているので、 $JIF$ は 2001 年と 2002 年の値の平均をとった。

表 5-10 各分野における雑誌の  $D$  平均値とインパクトファクター( $JIF$ )の間の Pearson 相関係数 ( $r$ )

分野	雑誌	$D$ mean	$JIF$	$r$
物性	European Physical Journal B	0.411	1.78	-0.872
	Journal of Physics - Condensed Matter	0.422	1.69	
	Physica B	0.444	0.64	
	Physical Review B	0.408	3.20	
無機	Inorganic Chemistry	0.453	2.95	-0.873
	Inorganica Chimica Acta	0.460	1.48	
	Journal of the Chemical Society - Dalton Transactions	0.434	2.92	
	Transition Metal Chemistry	0.476	0.86	
電気	IEE Proceedings - Circuits, Devices and Systems	0.437	0.24	0.185
	IEEE Transactions on Circuits and Systems I	0.488	0.80	
	IEEE Transactions on Microwave Theory and Techniques	0.449	1.39	
	Signal Processing	0.459	0.51	
生化	European Journal of Biochemistry	0.443	2.92	-0.298
	Journal of Biochemistry (Tokyo)	0.428	1.93	
	Journal of Biological Chemistry	0.430	6.98	
	Journal of Molecular Biology	0.430	5.59	
生理	Journal of General Physiology	0.432	5.34	-0.487
	Japanese Journal of Physiology	0.459	1.10	
	Journal of Physiology - London	0.444	4.56	
	Pflugers Archive European Journal of Physiology	0.432	1.66	
消化器	American Journal of Gastroenterology	0.468	3.75	-0.998
	Gastroenterology	0.446	13.23	
	Gut	0.463	6.25	
	Journal of Gastroenterology	0.473	1.35	

## 第5章 論文の引用持続性を示す指標の性質

この表に示すように、 $D$ と  $JIF$ の Pearson 相関係数は、電気分野を除くすべての分野で負であり、その相関はかなり強い。雑誌数が4と少ないので有意性を確認することはできないが、これは偶然とは言えないようである。このことについては5.3.5.2でも言及する。

### 5.3.5 引用持続性と他の論文特性との関係に関する線形重回帰分析

5.2.5で述べたように、重回帰分析は、 $C(T) \geq 5$ の論文から抽出した比較的小さい標本（各分野の標本サイズは表5-3を参照）に対して行った。

#### 5.3.5.1 予備的な分析—変数間の相関

重回帰分析に先立って、目的変数  $D$ と各説明変数との Pearson 相関係数  $r$ を求めた。その結果を表5-11に示す。 $C(T)$ は6分野すべて、*Tables*は5分野で  $r > 0.1$ である。一方、*Price*は4分野、*Figures*は3分野で  $r < -0.1$ である（用いたサンプルサイズでは、 $|r| > 0.1$ はほぼ10%有意に対応する）。従って、これらの4変数は、重回帰分析において  $D$ と有意な関係を示すことが示唆される。但し、 $|r|$ が0.3を超える場合はほとんどなく、それほど強い相関ではない。図5-8に、これら4変数に対して各分野で得られた  $r$ 値の95%信頼区間を示す。

表5-11  $D$ と各説明変数との Pearson 相関係数 ( $r$ )

説明変数	物性	無機	電気	生化	生理	消化器
<i>Authors</i>	-0.019	-0.092	-0.001	0.027	0.020	<b>0.106</b>
<i>Insts</i>	-0.027	<b>-0.143</b>	0.042	-0.009	0.025	0.073
<i>Countries</i>	-0.059	-0.085	0.036	0.023	-0.025	<b>0.171</b>
<i>Refs</i>	-0.009	-0.097	0.016	0.019	-0.068	<b>-0.101</b>
<i>Price</i>	<b>-0.102</b>	<b>-0.179</b>	-0.037	<b>-0.134</b>	-0.082	<b>-0.325</b>
<i>Length</i>	0.062	-0.045	0.043	0.030	-0.087	-0.028
<i>Figures</i>	0.006	-0.043	<b>-0.119</b>	<b>-0.115</b>	<b>-0.109</b>	-0.025
<i>Tables</i>	<b>0.105</b>	<b>0.152</b>	<b>0.128</b>	<b>0.126</b>	0.018	<b>0.284</b>
<i>Eqs</i>	0.057	0.004	0.028	<b>0.101</b>	<b>-0.164</b>	-
<i>Age</i>	-0.036	<b>-0.123</b>	0.007	0.034	-0.074	0.001
<i>RatePubl</i>	<b>0.163</b>	-0.024	0.082	-0.026	-0.005	<b>-0.101</b>
<i>MedCite</i>	-0.056	-0.067	<b>-0.173</b>	0.033	0.062	-0.064
$C(T)$	<b>0.222</b>	<b>0.104</b>	<b>0.300</b>	<b>0.198</b>	<b>0.185</b>	<b>0.287</b>

ボールド体： $r > 0.1$  ボールド斜体： $r < -0.1$

第5章 論文の引用持続性を示す指標の性質

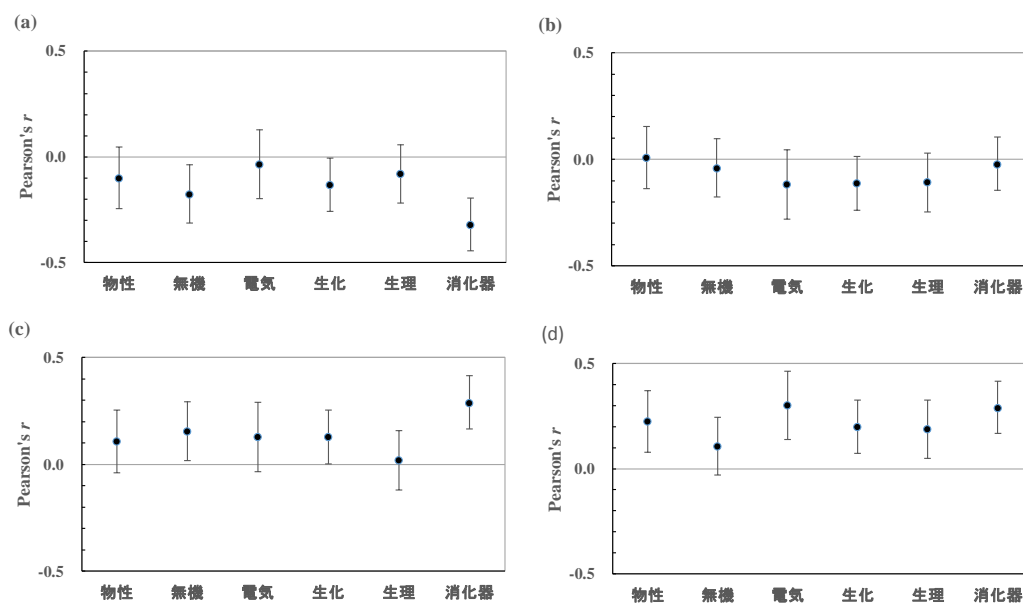


図 5-8  $D$  と各説明変数との間の Pearson 相関係数  $r$  (バーは 95%信頼区間を示す)

(a) Price, (b) Figures, (c) Tables, (d)  $\log[C(T)+1]$

説明変数相互間の相関の概要を表 5-12 に示す。この表には、Pearson 相関係数  $r$  の絶対値が 0.15 より大きい (概ね 5%有意に対応) 分野の数を示しており、右上が正の相関の数、左下が負の相関の数である。*Authors-Insts-Countries*、*Refs-Length-Figures*、*Age-RatePubl-MedCite* が、それぞれ相関が比較的高いグループを形成している。また、 $C(T)$  はいくつかの変数と有意な正の相関がある。3 つ以上の分野で  $r$  (の絶対値) の値が 0.5 を越える組み合わせは、*Authors-Insts*、*Insts-Countries*、*Refs-Length*、*Length-Figures* であった。しかしその値が 0.7 を越す例はほとんどなかったため、多重共線性の問題が生ずる可能性は低いと考え、全変数を重回帰分析に含め、ステップワイズ法により有意な変数を選択することにした。



表 5-12 説明変数間の相関

左上三角形部分は  $r > 0.15$ 、右下三角形部分は  $r < -0.15$  となる分野の数を示す。

	<i>Authors</i>	<i>Insts</i>	<i>Countries</i>	<i>Price</i>	<i>Refs</i>	<i>Length</i>	<i>Figures</i>	<i>Tables</i>	<i>Eqs</i>	<i>Age</i>	<i>RatePubl</i>	<i>MedCite</i>	<i>C(T)</i>
<i>Authors</i>		6	6	3	1	2	2	2		1			2
<i>Insts</i>			6	1	1	2	1	1		2		1	3
<i>Countries</i>					1	1						1	1
<i>Price</i>							1				1	2	6
<i>Refs</i>	1					6	5	2	3		1	1	5
<i>Length</i>	1	1					6	4	5		1	1	4
<i>Figures</i>		1						1	2				2
<i>Tables</i>				2			1		1				1
<i>Eqs</i>	2			1									
<i>Age</i>				3	1							6	6
<i>RatePubl</i>							1						3
<i>MedCite</i>								1					
<i>C(T)</i>			1							1			

5.3.5.2 重回帰分析

$D$ を目的変数、表5-2に示す諸特性を説明変数とする重回帰分析の結果は、6つの分野すべてにおいて有意であった（有意確率  $p$  は  $10^{-4}$  未満）。決定係数( $R^2$ )、自由度調整済み決定係数( $R_c^2$ )、分散比( $F$ )を表5-13に示す。 $R_c^2$ は0.1~0.3程度で高いとは言えない。

表5-13 重回帰分析の結果：決定係数( $R^2$ )、自由度調整済み決定係数( $R_c^2$ )、及び分散比( $F$ )。

	物性	無機	電気	生化	生理	消化器
$R^2$	0.136	0.131	0.177	0.153	0.130	0.350
$R_c^2$	0.116	0.108	0.154	0.138	0.112	0.330
$F$	6.76	5.61	7.54	10.20	7.23	17.23

表5-14は、目的変数  $D$  との関係が有意であった説明変数の標準化回帰係数( $\beta$ )の推定値と95%信頼区間である。有意な変数を選択する条件は、ステップワイズ法で投入上限10%、除去下限20%とした。つまり、この表には10%水準で有意とされた変数が含まれているため、95%信頼区間に0を含む変数も若干存在する。標本サイズがあまり大きくない(200前後)ため、有意水準をややゆるめに設定した。なお、消化器分野ではすべての論文において  $Eqs$  (数式の数) が0であったので、この変数は重回帰分析に含まれていない。

3つ以上の分野で選択されたのは次の5変数である。

(1) *Price*

電気を除く5分野で、 $D$ に対して負の効果が見られた。すなわち、参考文献中近年のもの(引用年齢が5年以内)の割合が高い論文は早期に引用される傾向がある。

(2) *Figures* と *Tables*

それぞれ3つの分野で  $D$  との関係が有意であるが、その関係は *Figures* では負、*Tables* では正である。つまり、図が多い論文は比較的早期に、表が多い論文は比較的遅延的に引用される。但し、これらの関係はそれほど強いものではない。

(3)  $\alpha(T)$  と  $JIF$

$\alpha(T)$ は全分野で比較的強い正の関係を、 $JIF$ は3つの分野でやや弱い負の関係を示す。これは、 $\alpha(T) \geq 5$ の全論文に対して行った相関分析の結果(5.3.3、5.3.4参照)を裏付けるものである。

表 5-14 重回帰分析の結果：標準化偏回帰係数( $\beta$ )<sup>a, b</sup>

説明変数	物性	無機	電気	生化	生理	消化器
<i>Authors</i>						
<i>Insts</i>		-.164±.139				
<i>Countries</i>						.136±.122
<i>Refs</i>						-.163±.132
<i>Price</i>	-.184±.149	-.223±.140		-.302±.143	-.232±.145	-.380±.123
<i>Length</i>		-.185±.171				
<i>Figures</i>			-.128±.152	-.136±.125	-.207±.157	
<i>Tables</i>		.226±.154	.185±.153			.111±.125
<i>Eqs</i>					-.136±.143	
<i>Age</i>						
<i>RatePubl</i>	.160±.142					
<i>MedCite</i>			-.212±.152			
<i>C(T)</i>	.308±.152	.241±.153	.340±.154	.423±.148	.354±.156	.464±.150
<i>JIF</i>	-.183±.147			-.144±.141		-.191±.154

a) ステップワイズ法で選択された変数に対し、推定値と 95%信頼区間を示す。

b) 斜体：95%信頼区間が 0 をまたぐ。

重回帰分析の結果は、前もって予備的に行った相関分析の結果と合致し、それを更に明確化するものであった。相関分析ではやや不明確であった  $D$  と  $C(T)$ 、 $Price$ 、 $Tables$ 、 $Figures$  との関係が、重回帰分析により明確になった。これらの変数では、 $D$  との関係が有意でない分野でも  $B$  の符号に矛盾はなかった。

上記の 5 つ以外の変数については、 $D$  との有意な関係は見られなかった。

各分野の標本論文は、それぞれ 4 つの雑誌からほぼ同数ずつ採られている。上記の重回帰分析の結果に雑誌のバイアスがかかっていないことを確認するため、24 の雑誌ごとに同じ方法で重回帰分析を行ってみた。標本サイズが小さくなるので全般に説明力は減ったが、 $C(T)$ 、 $Price$ 、 $Figures$ 、 $Tables$  はそれぞれ 15、7、2、5 の雑誌で  $D$  との関係が有意であり、分野別の分析と矛盾する結果は見られなかった。

## 5.4 考察

### 5.4.1 通算被引用数及びインパクトファクターと $D$ との関係

本章に述べた研究では次の結果が得られた。

- (1)  $D$  は  $C(T)$  とは正の相関がある (表 5-8 参照)。重回帰分析でも  $C(T)$  の偏回帰係数はすべての分野で有意な正值である (表 5-14 参照)。
- (2) しかし、 $D$  と  $C(T)$  の関係は線形ではなく、ある  $D$  で  $C(T)$  の平均値は最大値をとる (表 5-9、図 5-7 参照)。

(3) 一方、インパクトファクター(*JIF*)の高い雑誌では  $D$  はむしろ低くなる傾向がある(表 5-10 参照)。*JIF* の偏回帰係数は 3 つの分野で有意な負値であった(表 5-14 参照)。

(1)に関して、高被引用論文の引用が遅延的になることは、これまでのいくつかの研究でも報告されている(Line, 1984; Aversa, 1985; Levitt and Thelwall, 2008, 2009a; Costas et al., 2010; Wang, 2013)。しかし、他の論文特性を制御した重回帰分析によってこの関係を確認したのは、本研究が初めてである。

(2)に関しては、Costas et al. (2010)が、早期引用型[II]、遅延引用型[III]、標準型[IV]の 3 クラス間で論文の分野規格化被引用数(*CPPFCSm*)を比較している。被引用ウィンドウを発表後 5 年間、10 年間、20 年間と変えるに連れて、*CPPFCSm* の平均値は、 $IV > II > III \rightarrow IV > III > II \rightarrow III > IV > II$  と変動した。被引用ウィンドウ期間が 15 年である本章での結果では、 $D$  とともに  $\alpha(D)$  は増加する傾向を持つが、ある  $D$  で  $\alpha(D)$  の平均値は最大値をとる。これは Costas et al. の被引用ウィンドウ期間 10 年の傾向と一致する。

(3)についてはこれまで報告されていない。このことは、それぞれの雑誌の中では高被引用論文の方が高い引用持続性を持つ傾向があるが、雑誌間で比べると、*JIF* の高い雑誌には引用持続性の低い論文が含まれている割合が高いことを意味する。このやや矛盾した結果は、*JIF* が論文発表後 2 年以内の短期的引用に基づいていることと関係あるのではないかと一見考えられる。しかし、 $D$  の雑誌平均値と長期間(発表から 15 年間)の被引用数である  $\alpha(D)$  の雑誌平均値の間にも、次のように電気以外の 5 分野でかなり強い負の相関が見られた。

物性  $r = -0.812$   
 無機  $r = -0.756$   
 電気  $r = 0.670$   
 生化  $r = -0.248$   
 生理  $r = -0.499$   
 消化器  $r = -0.975$

6 分野中 1 分野に例外があるので決定的なことは言えないが、この結果は、短期的か長期的かによらず、引用インパクトの高い雑誌には引用持続性の低い論文が含まれる割合が高いという主張を支持する。図 5-9 は消化器分野でのこの状況を説明するものである。Gastroenterology、Gut、Am J Gastroenterol、J Gastroenterol の順に、グラフの右下から左上へ(すなわち  $\alpha(D)$  が高い方から低い方へ、 $D$  が低い方から高い方へ)分布の中心が移動している。しかし、全体としては(そしてそれぞれの雑誌内においても)  $\alpha(D)$  が高くなると  $D$  も高くなる右上がりの傾向が見られる。

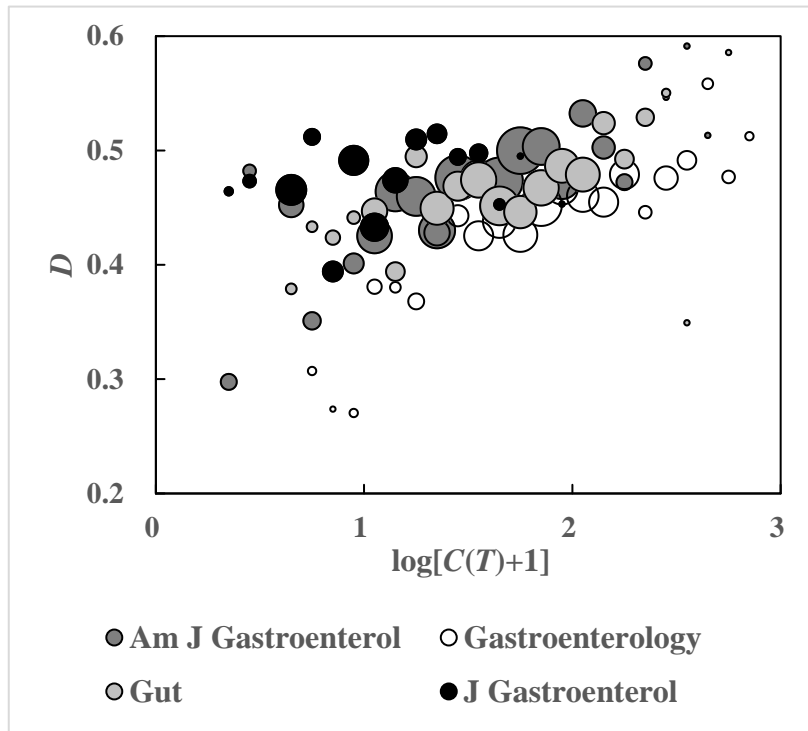


図5-9 消化器分野の4雑誌の間の  $\alpha(T)$ - $D$ 分布の比較

各バブルは、それぞれの  $\log[C(T)+1]$  区間における  $D$  の平均値。  
バブルの大きさは論文数に対応。

#### 5.4.2 $D$ と他の論文特性の間の関係

5.1 で述べたように、引用持続性と論文の特性の関係を論じた研究は少なく、van Dalen and Henkens (2005)、Costas et al. (2010)、Wang et al. (2015) を見る程度である。ここでは、これらの結果と本章での研究の結果を比較する。

van Dalen and Henkens (2005) は、1990–92 年の間に 17 の人口学雑誌に発表された 1,371 論文を、発表後 10 年間に得た引用のパターンによって [I] 引用僅少型、[II] 早期引用型、[III] 遅延引用型、[IV] 標準型の 4 クラスに分類した。そして、論文の諸特性が各クラスへの所属に与える影響を、クラス I を基準としたロジット多項回帰分析により検討した。説明変数のうち、発表誌がトップレベルであること、著者の名声の高さ（過去の被引用数が大）、論文長が、クラス I よりもクラス II、III、IV に属する確率を上昇させるが、著者数、著者所属国などは有意な効果を示さなかった。彼らはクラス II、III、IV 間の相互比較には言及していないので、本章の結果と直接の比較はできないが、結果を示すデータからは、有意であった 3 変数はクラス IV との相関が最も強いことが示唆される。クラス II と III の間では、トップジャーナルであることがクラス II の方に有利に働くが、他の 2 変数には差がない。

Costas et al. (2010)は、同様のクラス I、II、III、IVへの4分類（但し分類の基準は異なる）を用いて、WoSに収録された1983–2003年の期間の論文を2008年までの引用データにより分類し、クラスII、III、IVの3クラスの特性を比較した。クラスIIIはクラスIIに比べ著者数、著者所属機関数、所属国数、参考文献数が少なく、ページ数が多く、いずれもその差は有意であった。

Wang et al. (2015)は、2001年発表の論文が得た13年間の引用データから、Citation Delay（ここで言う  $D$ ）を従属変数とし、論文の学際性を表す因子（参考文献の論文が属する分野の分布に基づく）といくつかの制御変数を独立変数とする重回帰分析を行った。彼らの主眼は学際性と  $D$  の関係であるが、ここではより一般的な論文特性である制御変数の方に着目する。Wang et al.は、著者数、国際共著であること、参考文献数は  $D$  に対して有意な負の効果を、ページ数は有意な正の効果を持つという結果を得た。

以上の結果と本章の研究で得られた結果を対比したのが表5-15である。van Dalen and Henkensの結果は本章の重回帰分析の結果と矛盾しない。すなわち、*Authors*、*Length*、*MedCite*（それぞれ van Dalen and Henkens の著者数、論文長、著者の名声の高さに対応）は引用持続性とは明確な関係が認められないが、*JIF*（トップジャーナルに対応）は負の関係を示した。一方、本章の研究では引用持続性との明確な関係が認められなかった共同研究の拡がりの程度（著者数、著者所属機関数、所属国数）、参考文献数、論文の長さに対して、Costas et al.と Wang et al.は有意な関係を認めた。この理由として、以下のことが考えられる。

表 5-15 引用持続性と種々の論文属性の間の相関に関する先行研究と本章の研究の比較

属性	van Dalen & Henkens (2005)	Costas et al. (2010)	Wang et al. (2015)	本研究
著者数	なし	負に有意	負に有意	なし
所属機関数		負に有意		なし
所属国数		負に有意	負に有意	なし
参考文献数		負に有意	負に有意	なし
論文長	なし	正に有意	正に有意	なし
著者の業績	なし			なし
雑誌の名声	負に有意			負に有意
通算被引用数		正に有意		正に有意
Price指数				負に有意
図の数				負に有意
表の数				正に有意

- (1) それぞれ標本サイズが大きく異なる。本章の研究と van Dalen and Henkens の標本サイズは  $10^2 \sim 10^3$  程度であるのに対し、Costas et al.と Wang et al.の標本は極めて大きい（Costas et al.は  $n \approx 8 \times 10^6$ 、Wang et al.は  $n \approx 3 \times 10^5$ ）。従っ

て有意となる効果量に大きな相違がある。実際、Costas et al.に示されているデータでは、3クラス間の各変数の平均値の差はそれほど大きくない。また、Wang et al.における偏回帰係数の値から推定される独立変数の従属変数( $D$ )に対する効果もかなり小さい。Wang et al.による重回帰の $R^2$ は0.02程度である(本章の研究では0.13以上)。

- (2) 本章における重回帰分析では、通算被引用数、Price 指数、図や表の数が  $D$  と有意な関係を示した。これらの変数は、Wang et al.の重回帰分析には含まれていない。また、Costas et al.では引用持続性とそれぞれの変数の関係を個々に比較している。従って、彼らが得た引用持続性と共同研究の拡がりの程度、参考文献数、論文の長さの間の有意な関係は、交絡であった可能性がある。

(2)に関して、本章の重回帰分析で用いた変数の制御により共同研究の拡がりの程度、参考文献数、論文の長さとの関係が抑えられたという可能性を検討するため、 $D$ と最も関係が強い変数である  $\alpha(T)$ を除いて重回帰分析を行った。予想に反して、 $\alpha(T)$ を除くことにより有意性の向上あるいは効果量(標準化回帰係数)の増大があったケースよりも、その逆のケースがずっと多かった。*Authors*、*Insts*、*Refs*、*Length*、はやはりほとんどの分野で有意ではなく、一方、*Price*、*Figure*、*JIF*の各変数は、10%有意として選ばれた分野の数がそれぞれ5→3、3→1、3→1と減少した。つまり、これらの変数が持つ  $D$ との負の相関が、 $D$ と正の関係を持つ  $\log[\alpha(T)+1]$ の制御によってより明確に示されたということである。

### 5.4.3 将来の課題

5.3.2.3 で述べたように、 $D$ の分布における分野間の差はそれほど大きくないが、それでもある程度の差があることが判った(図 5-5 参照)。従って、被引用数そのものと同様、異なる分野の論文の  $D$ を直接比較することはできない。このためには、何らかの規格化が必要であるが、この問題は本研究の目的の範囲を超えている。

これまでの研究で、引用の持続性のパターンから論文を *sleeping beauty* (delayed)、*flash in the pan* (early-matured)、*normal* 等に分類する試みがいろいろ見られる。 $D$ は(他の測度とともに)この種の分類に利用することができるが、 $D$ の値が直接このような分類に結びつくものではない。引用持続性に関する別の指標である“Beauty Coefficient”を最近提案した Ke et al. (2015)は、この指標の分布に関して“there are no clear demarcation values that allow us to separate sleeping beauties from “normal” papers: delayed recognition occurs on a wide and continuous range ...”と述べている。 $D$ についても同じことが言える。定量的指標と定性的カテゴリーの関係づけは将来の課題である。

同じ論文でも、観測期間によって  $D$ は変化する。その変化は論文間で一様ではな

いであろうから、論文発表から 15 年間の引用履歴から得られた  $D$  の分布形やその他の性質は、この期間をもっと長くしたとき変化があるかもしれない。また、 $D$  による論文のランキングも変わるかもしれない。これも今後の検討課題である。

論文の通算被引用数が大きく違っても、引用履歴のパターンが同じであれば  $D$  は同じ値になる。たとえば論文 A に比べて論文 B の毎年の被引用数が常に 2 倍であれば、両者の  $D$  は等しい。しかしその逆は成り立たず、2 つの論文の  $D$  の値が非常に近くても引用履歴のパターンが大きく違うことがある。最近 Sun et al. (2016) は、このようなパターンの違いをある程度識別できる指標として“obsolescence vector”を導入した。この指標は、Li et al. (2014) が提案したパラメータ  $G_s$  (本研究の  $D$  に似た測度) と累積引用曲線の drastic fluctuation を検出するパラメータ  $A$  の 2 成分から成るベクトルである。引用履歴の微細な構造の違いを識別する持続性指標の開発も、今後の研究課題である。

## 5.5 本章の結論

5.1 で示した 3 つの研究目的に沿って、本章の研究で得られた主な知見をまとめる。

**研究目的 1** : 引用持続性指標 Citation Delay ( $D$ ) の分布の特徴を検討する。

$D$  の主な特徴は次の通りである。

- ・論文の引用計測期間全域の情報を反映している。
- ・0 から 1 の間の範囲を取り、引用が遅延型になるほど 1 に近づく。
- ・ほぼ正規分布に従う。ある程度以上の被引用数の論文に限定するほど分布の幅が狭くなるとともに、正規分布との近似度が高まる。
- ・被引用数そのものに比べて分野間の格差は小さい。

以上はいずれも引用持続性の指標として好ましい性質である。

**研究目的 2** : 異なる分野において、 $D$  と論文の被引用数の関係を検討する。

被引用ウィンドウ 15 年間のデータに基づき、6 つの分野すべてにおいて  $D$  と  $\log[C(T)+1]$  の間に有意な正の相関があることを示した (Pearson 相関係数は 0.15 ~ 0.35)。すなわち、引用持続性が高い論文の方が通算被引用数が多い傾向がある。しかしこの関係は線形ではなく、 $D$  のある値で  $\log[C(T)+1]$  の平均値は最大となる。

**研究目的 3** : 異なる分野において、論文の諸特性と引用持続性の関係を検討する。

本研究は、引用持続性と他の論文特性の関係について体系的に分析した最初の研究であり、6 分野での重回帰分析の結果、いくつかの特性について有意な関係を見出した。

Price 指数は  $D$  と負の有意な関係がある。第 4 章の結果と合わせて考えると、参考文献中近年のもの比率が高い論文は長期的に高い引用を得るが、それは比較的



## 第5章 論文の引用持続性を示す指標の性質

早期になされる傾向がある。

図の数が多い論文は早期に引用され、表の数が多い論文は持続的に引用される傾向がある（それほど強い傾向ではないが）。

それぞれの雑誌の中では高被引用論文の方が高い引用持続性を持つ傾向があるが、雑誌間で比べると、インパクトの高い雑誌には引用持続性の低い論文が含まれている割合が高いという、一見矛盾した結果が得られた。この傾向は、雑誌インパクトに短期的測度（2年間インパクトファクター）を用いても、長期的測度（15年間の被引用数の雑誌平均）を用いても同じである。

## 第6章

### 結論

本研究では、6つの異なる主題分野（物性物理学、無機・核化学、電気・電子工学、生化学・分子生物学、生理学、消化器病学）において、同じ年（2000年）に発表された原著論文のサンプルを用いて、それら論文の長期的被引用数と引用持続性の性質を考究した。その中心は、長期的被引用数と引用持続性に及ぼす外在的要因（論文の質や内容に直接関係しない要因）の影響、及び長期的被引用数と引用持続性の間の相互関係の解明であった。

第1章で本研究の背景と目的・意義及び概要を述べ、第2章で先行研究をレビューした。第3章では、内在的要因と外在的要因の分類、及びこれらの要因の可操作化（測度による表現）について論じ、本研究で扱う要因と測度を選定した後、最も測定が難しい測度である著者の過去論文数の正確な測定の方法を提示した。

第4章では、発表後6-7年間及び11-12年間に得た被引用数が、どのような外在的要因にどの程度影響されるかを体系的に検討した。上記の6つの分野についてこれを行うことにより、影響を与える主要因について、ある程度分野共通の傾向があることを見出した。

第5章では、論文の引用持続性を測るのに最も適切と考えられる指標（Wang et al.により提案された Citation Delay）について、上記6分野の論文の引用履歴データを用いて、この指標の分布、及び長期的被引用数との関係を明らかにするとともに、第4章と同様の外在的要因との関係を分析した。

得られた結果とそれに基づく考察は各章に記しているが、本章では、まとめとして以下について記述する。

- (1) 今研究で得られた主要な知見（特に、長期的被引用数と引用持続性に共通に見られる傾向、及び両者の関係について）
- (2) 本研究のオリジナリティ
- (3) 本研究の限界
- (4) その他、若干の考察

#### 6.1 本研究で得られた主要な知見

##### 6.1.1 複数の分野に共通な傾向の確認

論文の長期被引用数（以降  $C$  と表記する）に影響を与える要因について、2.1 に示したように多くの研究がなされている。しかしながら、これらの研究の多くは、ある単一の要因に着目している（あるいは複数の要因をそれぞれ独立に見ている）ために、要因間の相互作用が明らかでない。いくつかの研究では、重回帰モデルを用いて様々な潜在要因を総合的に考慮しているが、対象の論文集合が特定の分野や雑誌に限定されているため、結論の一般性に限界がある。

一方、論文の引用持続性が持つ性質については、5.1 に述べたように、そもそも引用持続性の測度となる定量的指標が未確立であり、最も確と思われる指標である Wang et al. (2015) 提案の Citation Delay（本論文で  $D$  と表記）についても、その性質はほとんど研究されていない。

そこで本研究では、 $C$  及び  $D$  に関して、分野を越えた一般的傾向が見られるか否かを明らかにすることに着目し、本章冒頭に示した 6 つの異質な分野を選んで、それらを比較検討できるよう系統的に分析を行った。

種々の論文特性が  $C$  に及ぼす影響については負の 2 項重回帰(NBMR)モデルにより(第 4 章)、 $D$  に及ぼす影響については線形重回帰モデルにより(第 5 章)分析した。このいずれについても、6 つの分野の間には、選ばれた予測変数とその有意性の大きさにおいてある程度の共通性があった。この結果についてはそれぞれの章でも述べているが、6.1.3 に両者の結果を比較しながら考察する。

これらの重回帰モデルの当てはまりのよさ(goodness of fit)はそれほど高くはない。 $C$  の予測に対する NBMR モデルの調整済み擬決定係数(pseudo  $R^2$ )は分野により 0.25~0.5 程度、 $D$  の予測に対する重回帰モデルの自由度調整済み決定係数( $R^2$ )は分野により 0.1~0.3 程度である。今回用いた説明変数は全て、論文の質や内容に直接関係するものではない「外在的要因」であることを考えれば、これは予想される結果である。本研究の目的は、高い適合度のモデルを求めることではなく、これら外在的要因に基づく期待値のベースラインとなるモデルを探ることである。

### 6.1.2 $C$ と $D$ の間の関係

第 5 章で得られた結果を要約すると、次のようになる。

- (1)  $D$  は  $C$  とは正の相関がある。 $D$  を目的変数とする重回帰分析でも、 $C$  はすべての分野で有意な正の予測変数である。しかし、 $D$  と  $C$  の関係は線形ではなく、ある  $D$  で  $C$  の平均値は最大値をとる。
- (2) 一方、雑誌単位で見ると、引用インパクトの高い雑誌では  $D$  はむしろ低くなる傾向がある。このことは、雑誌の引用インパクトにインパクトファクターのような短期の指標を用いても、発表後 15 年間の平均被引用数のような長期の指標を用いても同様であった。

つまり、それぞれの雑誌の中では高被引用論文の方が高い引用持続性を持つ傾向があるが、雑誌間で比べると、インパクトの高い（短期的か長期的かによらず）雑誌には引用持続性の低い論文が含まれている割合が高い。

これは興味ある結果であるが、(2)については、各分野で4つの雑誌のみの比較に留まること、6分野中1分野の例外があったことから、強い結論とは言えない。いくつかの分野で、インパクトの異なる多くの雑誌を含めた分析が必要である。

### 6.1.3 $C$ 及び $D$ に影響を及ぼす特性の比較

第4章のNBMR分析と第5章の重回帰分析で用いた論文サンプルはほぼ同じであり（後者では $C < 5$ の論文を除外）、また、用いた説明変数も、後者において $C$ と $JIF$ を加えた以外は共通である（前者では $JIF$ の代わりに雑誌ダミー変数を用いた）。そこで、これらの結果に基づき、 $C$ との関係が強い特性と $D$ との関係が強い特性を比較する。

$C$ と関係の強い説明変数は、第一にPrice指数( $Price$ )、次に参考文献数( $Ref$ )、やや弱いのが著者数( $Authors$ )と第一著者過去論文の年あたり被引用数中央値( $MedCite$ )で、いずれも正の関係であった(4.3.2.3及び表4-7参照)。一方、 $D$ との関係が比較的強かったのは $Price$ 、図の数( $Figures$ )、表の数( $Tables$ )で、 $Price$ と $Figures$ は $D$ と負の関係、 $Tables$ は正の関係であった(5.3.5.2及び表5-14参照)。

$C$ と $D$ のどちらも強い関係を示したのは $Price$ であるが、その符号の向きは逆である。すなわち、 $Price$ の値が高い（参考文献中近年のもの比率が高い）論文は長期的に高い引用を得るが、それは比較的早期になされる傾向がある。 $Price$ の意味から考えるとこの結果は肯かれるが、全体的には通算被引用数が多い論文には遅延引用型の方が多い(5.3.3参照)ので、 $Price$ の高い論文はその中でやや特異な存在であると言える。

$Price$ 以外で $D$ と有意な関係が見られた $Figures$ と $Tables$ は、 $C$ との関係は明らかでなかった。論文の被引用数を説明する重回帰分析の説明変数に図の数や表の数を含めた報告はこの他にも少数存在する(Snizek et al., 1991; Haslam et al., 2008)が、いずれも有意な関係は示されていない。本研究では、被引用数との関係は否定的である図の数や表の数が、引用持続性と関係を持つという結果を初めて示した。図の数が多い論文は早期に引用され、表の数が多い論文は持続的に引用される傾向がある（それほど強い傾向ではないが）という結果は、早期に注目を得たい場合は図を多くし、じっくりと内容の検討を求めたい場合は表を多くするという研究者の行動を示唆するものかもしれない。

### 6.1.4 引用持続性指標の性質の分析

引用持続性指標がどのような性質を持つかについては、本研究で用いた  $D$  を含め、ほとんど解明されておらず、これを複数の分野において検討した本研究の意義は大きいと考えられる。6.1.2、6.1.3 に述べたことのほか、 $D$  の分布について分析を行い、どの分野においても正規分布に非常に近い分布を示すこと、 $C$  の分布に比べて分野間の差がかなり小さいことを初めて確認した(5.3.2 参照)。

## 6.2 本研究のオリジナリティ

ここでは、6.1 の記述と重複するところがあるが、1.2 で本研究のオリジナルな点としてあげた事項が実際に達成されたかどうかという観点からまとめる。

### 6.2.1 研究全体にわたってのオリジナリティ

#### (1) 複数の分野への同様の分析方法の適用による一般的結論の導出

論文の被引用数に及ぼす外在的要因の影響(第4章)、及び論文の引用持続性の性質(第5章)に関して、6つの異なる分野において同様の方法で分析を試み、ある程度分野に共通する傾向を示したことが、本研究の最大の特徴である。2.1.3 で述べたように、論文の被引用数に対する影響要因を総合的に考慮した先行研究は少なくないが、これらは特定の分野、特定の雑誌、あるいは特定の国の論文に対象を限定しており、得られた結論が一般化できるとは必ずしも言えない。本研究で取り上げた6つの分野は、Web of Science (以下 WoS と略記) の約 200 に及ぶ自然科学主題カテゴリーの一部ではあるが、理工学系(物性物理学、無機・核化学、電気・電子工学)とライフサイエンス系(生化学・分子生物学、生理学、消化器病学)、基礎系(物性物理学、無機・核化学、生化学・分子生物学、生理学)と応用系(電気・電子工学、消化器病学)にわたっており、これらについて共通な傾向が見られたことから、ある程度一般的な結論を得たと考えられる。以下に述べるオリジナリティはすべて、この6分野共通の傾向として明らかにされたものである。

#### (2) 論文の被引用数に影響を及ぼす外在的要因の体系的分類とその適用

論文の被引用数に影響を及ぼす可能性のある外在的要因を、これまでの研究を整理することにより9つに分類した(第3章)。第4章ではこのうちの5種の分類からそれぞれ2種以上の測度を選び、これらの測度を説明変数として、論文の被引用数への影響について検討した。別の要因である引用計測期間の影響については、他の要因とは異なるアプローチが必要であるので、別途に第5章で検討した。要因を3つ程度に分類した研究はこれまでもあるが(2.1.3.2を参照)、それらに比べ本研究での分類はより体系的と

考えられる。

### 6.2.2 第3章の研究についてのオリジナリティ

#### (1) 比較的簡便な著者識別方法の提案

ここで提案した著者識別法（同名著者の論文を、ターゲットとする著者のものか同名の別著者のものかを判別）は、より厳密な方法に比べいくつかの限界があるが、大量のデータを比較的少ないステップで処理できるという利点があり、同名異人著者による「偽」論文が混入した論文集合から「真」論文を95%程度の精度で識別することができる。

#### (2) 著者識別に有効な論文属性の実証

著者識別に用いた論文の属性のうち以下のものについては、その有効性を初めて実証した：①直接引用関係または共引用関係の存在（確実に同著者の論文であると言える）、②特定の国への著者所属、③比較する論文の間の発表年の差。

### 6.2.3 第4章の研究についてのオリジナリティ

#### (1) 長期の被引用数の予測における結果の安定性

被引用数を予測する負の2項重回帰(NBMR)分析では、目的変数に、比較的長期の被引用数を2通り— $C6$ （論文発表後6-7年間）と $C11$ （論文発表後11-12年間）—設定した。2つの目的変数の結果がほぼ同様であったことは、得られた結果の安定性を示唆するものである。また、論文発表後11-12年間という期間は、これまでの類似の研究の中で最長であり、この点でも信頼性の高い結果を得たと考えられる。表4-8に示した総合的重回帰研究のうち、Haslam et al. (2008)とFu and Aliferis (2010)は9-10年、他は6年以下である。

#### (2) 論文の被引用数に強い影響を及ぼす要因の発見

MBMR分析では、Price指数が全ての分野で最も強力な被引用数への影響要因であり、次いで参考文献数が重要であることが見出された。参考文献数は多くの先行研究で有意な関係が示されているが、Price指数の予測変数としての重要性を全分野で示したことは、本研究の主要な発見の一つである。

### 6.2.4 第5章の研究についてのオリジナリティ

#### (1) 引用持続性指標の性質の解明

これまで提案されたものの中で最適の引用持続性測度と考えられる Citation Delay ( $D$ ) の性質を初めて詳しく検討し、以下の結果を得た。

- ・ $D$ の分布はかなり正規分布に近く、ある程度以上の被引用数の論文に限定するほ

どその標準偏差は小さくなる。また、被引用数そのものに比べて分布の分野間格差が小さい。

- $D$ と長期被引用数の間には正の相関があるがその関係は線形ではなく、ある  $D$  で長期被引用数の平均値は最大値をとる。しかし、雑誌単位で見ると、引用インパクトの高い雑誌では  $D$ はむしろ低くなるという傾向が見出された。

### (2) 被引用数と引用持続性に影響を与える要因の比較

第4章と第5章の結果を比較することにより、外在的要因の被引用数への影響と引用持続性への影響の違いが明らかになった。Price 指数の値が高い（参考文献中近年のもの比率が高い）論文は長期的に高い引用を得るが、それは比較的早期になされる傾向がある。また、図の数や表の数は被引用数との有意な関係は見られないが、図の数が多い論文は早期に引用され、表の数が多い論文は持続的に引用される傾向がある

## 6.3 本研究の限界

6.1.2において、雑誌インパクトと其中的論文の引用持続性の関係については、確定的な結論とは言えないことを述べた。この他に、本研究の限界と考えられることを以下に記す。

### 6.3.1 内在的要因と外在的要因の交絡の可能性

論文の被引用数及び引用持続性に及ぼす外在的要因の影響を検討する際、被引用数または引用持続性指標を目的変数、種々の外在的要因（を可操作化した測度）を説明変数とする重回帰分析法を用いた。このとき、内在的要因を表す測度は含めていない。このことは、内在的要因と外在的要因は独立（相関がない）と仮定していることを意味する。両者に相関がある場合、内在的要因が外在的要因の交絡因子となり、実は内在的要因の効果であるのに見かけ上外在的要因の影響として現れ、結果にバイアスをもたらす可能性がないとは言えない。但し、2.1.3に述べた先行研究のうち重回帰分析等の説明変数に内在的要因を含めたもの(Bornmann and Daniel, 2008; Haslam et al., 2008; Lokker et al., 2008; Chen, 2012 など)に、他と著しく結果が異なるものはないので、重要な問題にはならないと考えられる。

### 6.3.2 サンプル論文の発表年の限定について

本研究で用いたサンプル論文は、同じ年(2000年)に発表された論文に限定した。これは、被引用数その他の性質に発表年によるバイアスを生じさせないためである。しかし、2000

年に発表された論文のみに基づく本研究の結果を一般化することには、幾分問題があるかもしれない。結論をより一般化するには、別の年に発表された論文で同じ実験を行うことであろう。最近の論文では、ここで得たものとは異なる傾向が見出される可能性もある。しかし、引用インパクトや引用持続性を見るにはある程度長い期間が必要である。そうすると、最近の論文の分析は、もう少し時間が経ってから行うことが適当である。引用分析研究には、発表後相当期間を経た論文を対象にせざるを得ないという限界が伴うことに留意する必要がある。

### 6.3.3 著者実績測度に第一著者のデータのみを使用

重回帰分析における著者実績に関する説明変数(過去の論文数、それらが得た被引用数、活動期間)は、第一著者のデータを用いた。先行研究によっては、最もインパクトの高い著者のデータ、全著者の平均値等を用いて、これらと被引用数の間に有意な関係を見出しているものもある。本研究で、著者実績と被引用数の間に強い関係は見られなかった(効率実績指標の *RatePubl* と *MedCites* はいくつかの分野で中程度の有意な影響を示した)のは、第一著者のみのデータを用いたことによるかもしれない。

### 6.3.4 引用持続性の研究に15年間の引用ウィンドウのみを使用

引用持続性指標  $D$  の性質は、論文発表後15年間の引用ウィンドウによった。これは十分長期の期間であるが、当然、観測期間によって  $D$  は変化する。その変化は論文間で一様ではないであろうから、 $D$  の分布形やその他の性質は、この期間をもっと長くしたとき変化があるかもしれない。また、 $D$  による論文のランキングも変わるかもしれない。より長期の引用ウィンドウで分析を行い、これらの傾向に変化がないかどうか確認することが望ましい。

## 6.4 その他の考察

### 6.4.1 ピアレビューによる論文評価との関係

業績評価への計量書誌学データの利用の可能性を論ずる目的で、被評価対象(論文、研究者、研究グループ等)に対するピアレビューによる研究評価の結果と、それらの引用スコアの相関を調べた研究がいろいろ行われている(Oppenheim, 1997; Rinia et al., 1998; 林, 2003; Abramo et al, 2011b, 2013; Mryglod et al, 2013a, 2013b)。これらの相関はいずれも「ほどほど」(相関係数 0.3~0.6 程度)であるが、その他



にほぼ共通することがある。それは、ピアによる評価が低い被評価対象が多くの引用を得ることはまずないが、ピアによる評価が高ければ被引用度も高いとは限らないということである。図 6-1 はこの関係を模式的に示したものである。

図 5-7 を見ると、論文の  $D$  と被引用数の関係もややこれに似ている。これだけから結論できることではないが、このことは、被引用数そのものよりもその持続性の方が、論文の質（ピア評価の結果）とより密接に結びついているという可能性を示唆するものかもしれない。少なくとも、実際のピア評価のデータを使ってこの仮説を検証する価値はあると考えられる。ただ、引用持続性のデータは論文発表後相当長期間を経ないと取得できないので、仮にこの関係が立証できても、それを実際の研究評価に利用することは難しいであろう。

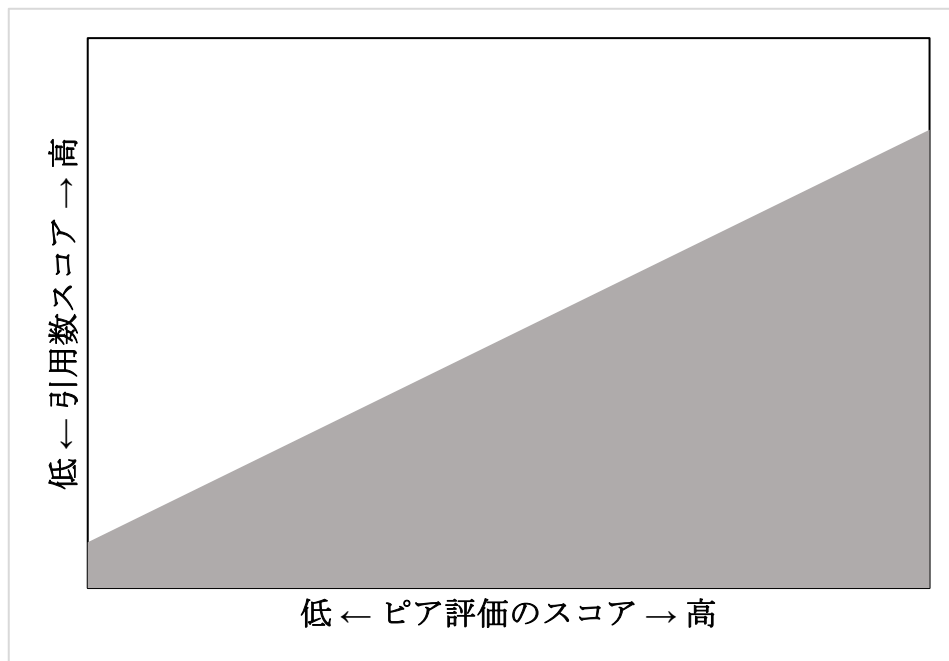


図 6-1 ピア評価のスコアと引用数スコアの関係を示す概念図

多くのデータは影を付けた部分に存在する。

#### 6.4.2 主題分野の設定について

本研究では、計量書誌分析でよく使われる WoS の主題カテゴリー (Subject Category) によって分野を設定した。しかし、今回得られた結果は、それらの分野を更にいくつかのサブフィールドに分割して分析する必要があることを示唆しているのかもしれない。Price 指数は、すべての分野で  $C$  と  $D$  のどちらに対しても強い予測変数であった。また、参考文献数は多くの分野で  $C$  に対する有意な影響要因となった。これら 4 つの変数 ( $C$ 、 $D$ 、Price 指数、参考文献数) はすべて引用行動に関する測度であり、サブフィールドごとに

引用行動が異なるとすれば、これらの変数間に強い相関が得られる可能性が高いからである。これに関して Moed (1989)は、同じ分野に属するサブフィールドの間で、参考文献数、Price 指数、及び被引用数の平均値が大きく異なる場合があることを示している。また、Ruiz-Castillo and Waltman (2015)は、引用の分野規格化を行うとき WoS の主題カテゴリーはなお不均質であると指摘し、論文の引用関係に基づき、数千の分野への分割が適当と結論した。WoS の主題カテゴリーより狭いサブフィールドによる引用行動の違いについて、更に検討が望ましい。

## 参 考 文 献

- Abramo, G., Cicero, T and D'Angelo, C. A. (2011a). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics*, 5(4), 659–667.
- Abramo, G., Cicero, T. and D'Angelo, C. A. (2013). National peer-review research assessment exercises for the hard sciences can be a complete waste of money: the Italian case. *Scientometrics*, 95(1), 311–324.
- Abramo, G., D'Angelo, C. A. and Di Costa, F. (2011b). National research assessment exercises: a comparison of peer review and bibliometrics rankings. *Scientometrics*, 89(3), 929–941.
- Akcan, D., Axelsson, S., Bergh, C., Davidson, T. and Rosen, M. (2013). Methodological quality in clinical trials and bibliometric indicators: no evidence of correlations. *Scientometrics*, 96(1), 297–303.
- Aksnes, D. W. (2003a). A macro study of self-citation. *Scientometrics*, 56(2), 235–246.
- Aksnes, D. W. (2003b). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159–170.
- Aksnes, D. W. (2006). Citation rates and perceptions of scientific contribution. *Journal of the American Society for Information Science and Technology*, 57(2), 169–185.
- Aksnes, D. W. (2008). When different persons have an identical author name. How frequent are homonyms? *Journal of the American Society for Information Science and Technology*, 59(5), 838–841.
- The American Society for Cell Biology. (2013). “The San Francisco Declaration on Research Assessment (DORA)”. <http://www.ascb.org/dora/> (参照 2018-02-12).
- Aversa, E.S. (1985). Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature. *Scientometrics*, 7(3/6), 383–389.
- Avramescu, A. (1979). Actuality and obsolescence of scientific literature. *Journal of the American Society for Information Science*, 30(5), 296–303.
- Basu A. and Lewison, G. (2005). Going beyond journal classification for evaluation of research outputs. *Aslib Proceedings*, 57(3), 232–246.
- Baumgartner, S. E. and Leydesdorff, L. (2014). Group-based trajectory modeling (GBTM) of citations in scholarly literature: Dynamic qualities of "Transient" and "Sticky Knowledge Claims". *Journal of the Association for Information Science and Technology*, 65(4), 797–811.
- Bookstein, A. and Yitzhaki, M. (1999). Own-language preference: A new measure of

- "Relative Language Self-Citation". *Scientometrics*, 46(2), 337–348.
- Bordons, M., Aparicio, J. and Costas, R. (2013). Heterogeneity of collaboration and its relationship with research impact in a biomedical field. *Scientometrics*, 96(2), 443–466.
- Bornmann, L. and Daniel, H.-D. (2006). Selecting scientific excellence through committee peer review ? A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics*, 68 (3), 427–440.
- Bornmann, L. and Daniel, H.-D. (2007). Multiple publication on a single research study: Does it pay? The influence of number of research articles on total citation counts in biomedicine. *Journal of the American Society for Information Science and Technology*, 58(8), 1100–1107.
- Bornmann, L. and Daniel, H.-D. (2008). Selecting manuscripts for a high-impact journal through peer review: A citation analysis of Communications that were accepted by *Angewandte Chemie International Edition*, or rejected but published elsewhere. *Journal of the American Society for Information Science*, 59(11), 1841–1852.
- Bornmann, L., Mutz, R. and Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden Ranking 2011/2012. *Journal of the American Society for Information Science and Technology*, 64(8), 1649–1658.
- Bornmann, L., Schier, H., Marx, W. and Daniel, H.-D. (2011). Is interactive open access publishing able to identify high-impact submissions? A study on the predictive validity of Atmospheric Chemistry and Physics by using percentile rank classes. *Journal of the American Society for Information Science and Technology*, 62 (1), 61–71.
- Burrell, Q. L. (2003). Age-specific citation rates and the Egghe–Rao function. *Information Processing and Management*, 39(5), 761–770.
- Burrell, Q. L. (2005). Are "Sleeping Beauties" to be expected? *Scientometrics*, 65(3), 381–389.
- Callaham, M., Wears, R. L. and Weber, E. (2002). Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA*, 287(21), 2847–2850.
- Cano, V. and Lind, N.C. (1991). Citation life cycles of ten citation classics. *Scientometrics*, 22(2), 297–312.
- Chen, C. (2012). Predictive effects of structural variation on citation counts. *Journal of the American Society for Information Science and Technology*, 63 (3), 431–449.
- Costas, R., van Leeuwen, T. N. and van Raan, A. F. J. (2010). Is scientific literature

- subject to a Sell-By-Date? A general methodology to analyze the durability of scientific documents. *Journal of the American Society for Information Science and Technology*, 61(2), 329–339.
- Cota, R. G. Ferreira, A. A., Nascimento, C., Gonçalves, M. and Laender, A. H. F. (2010), An unsupervised heuristic-based hierarchical method for name disambiguation in bibliographic citations. *Journal of the American Society for Information Science and Technology*, 61(9), 1853–1870.
- Cronin, B. and Shaw, D. (1999), Citation, funding acknowledgment and author nationality relationships in four information science journals. *Journal of Documentation*, 55(4), 402–408.
- Danell, R. (2011). Can the quality of scientific work be predicted using information on the author's track record? *Journal of the American Society for Information Science and Technology*, 62 (1), 50–60.
- Davis, P. M. (2009). Author-choice open-access publishing in the biological and medical literature: A citation analysis. *Journal of the American Society for Information Science and Technology*, 60(1), 3–8.
- Davis, P. M. and Fromerth, M. J. (2007). Does the arXive lead to higher citations and reduced publisher downloads for mathematics articles? *Scientometrics*, 71(2), 203–215.
- Davis, P. M., Lewenstein, B. V. Simon, D. H, Booth, J. G. and Connolly, M. J. L. (2008). Open access publishing, article downloads, and citations: randomised controlled trial. *BMJ*, 337, a568 (6p.).
- Della Briotta Parolo, P., Pan, R. K., Ghosh, R., Huberman, B. A., Kaski, K. and Fortunato, S. (2015). Attention decay in science. *Journal of Informetrics*, 9(4), 734–745.
- Didegah, F. and Thelwall, M. (2013). Determinants of research citation impact in nanoscience and nanotechnology. *Journal of the American Society for Information Science and Technology*, 64 (5), 1055–1064.
- Diodato, V. (1994). Dictionary of bibliometrics. New York, Haworth Press, Inc. (訳書 : ヴァージル・ディオダート (2008). 計量書誌学辞典. 芳鐘冬樹 ; 岸田和明 ; 小野寺夏生 訳. 東京, (社)日本図書館協会)
- Egghe, L. and Ravichandra Rao, I. K. (1992). Citation age data and the obsolescence function: Fits and explanations. *Information Processing and Management*, 28(2), 201–217.
- Fanelli, D. (2013). Positive results receive more citations, but only in some disciplines. *Scientometrics*, 94(2), 701–709.

- Figg, W. D., Dunn, L., Liewehr, D. J., Steinberg, S. M., Thurnman, P. W., Barrett, J. C. et al. (2006). Scientific collaboration results in higher citation rates of published articles. *Pharmacotherapy*, 26(6), 759–767.
- Fu, L. D. and Aliferis, C. F. (2010). Using content-based and bibliometric features for machine learning models to predict citation counts in the biomedical literature. *Scientometrics*, 85(1), 257–270.
- Garfield, E. (1980). Premature discovery or delayed recognition. Why? *Essays of an Information Scientist*, 4, 488–493.
- Garfield, E. (1984a). The 100 most-cited papers ever and how we select citation classics. *Current Contents*, (23), 3-9.
- Garfield, E. (1984b). 100 Classics from The Lancet. *Current Contents*, (39), 3-13.
- Gazni, A. and Didegah, F. (2011). Are the abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world. *Journal of Information Science*, 37(3), 273-281.
- Glänzel, W and Moed, H. F. (2013). Opinion paper: thoughts and facts on bibliometric indicators. *Scientometrics*, 96(1), 381–394.
- Glänzel, W. and Schoepflin, U. (1995). A bibliometric study on ageing and reception process of scientific literature. *Journal of Information Science*, 21(1), 37–53.
- Glänzel, W., Schlemmer, B. and Thijs, B. (2003). Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. *Scientometrics*, 58(3), 571–586.
- Glänzel, W. and Thijs, B. (2004). The influence of author self-citations on bibliometric macro indicators. *Scientometrics*, 59(3), 281–310.
- Glänzel, W., Thijs, B. and Schlemmer, B. (2004). A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*, 59(1), 63–77.
- Gumpenberger, C. Glänzel, W. and Gorraiz, J. (2016). The ecstasy and the agony of the altmetric score. *Scientometrics*, 108(2), 977–982.
- Han, H., Giles, C. L., Zha, H., Li, C. and Tsioutsoulouklis, K. (2004). “Two supervised learning approaches for name disambiguation in author citations”. Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Tucson, Arizona, USA, 2004-07-7/11, 296–305.
- Han, H., Zha, H. and Giles, C. L. (2005). “Name disambiguation in author citations using a K-way spectral clustering method”. Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Denver, Colorado, USA, 2005-07-7/11, 334–343.
- Haslam, N., Ban, L., Kaufmann, L., Loughnan, S., Peters, K., Whelan, J. and Wilson, S. (2008). What makes an article influential? Predicting impact in social and personality

- psychology. *Scientometrics*, 76(1), 169–185.
- 林隆之(2003). ビブリオメトリクスによるピアレビュー支援可能性の検討 理学系研究評価の事例分析から. 大学評価, (3), 169-187.
- He, Z.-L. (2009). International collaboration does not have greater epistemic authority. *Journal of the American Society for Information Science and Technology*, 60 (10), 2151–2164.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431.
- Hönekopp, J. and Khan, J. (2012). Future publication success in science is better predicted by traditional measures than by the h index. *Scientometrics*, 90(3), 843–853.
- Hsu, J.-W. and Huang, D.-W. (2011). Correlation between impact and collaboration. *Scientometrics*, 86 (2), 317–324.
- Huang, J., Ertekin, S. and Giles, C. L. (2006). “Efficient name disambiguation for large-scale databases”. *Proceedings of the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Berlin, Germany, 2006-10-19, 536–544.
- Ibáñez, A., Bielza, C. and Larranaga, P. (2013). Relationship among research collaboration, number of documents and number of citations: a case study in Spanish computer science production in 2000–2009. *Scientometrics*, 95(2), 689–716.
- Kanani, P. and McCallum, A. (2007). “Efficient strategies for improving partitioning-based author coreference by incorporating Web pages as graph nodes”. *Proceedings of the AAAI 6th International Workshop on Information Integration on the Web*, Vancouver, Canada, 2007-07-23, 38-43.
- Kanani, P., McCallum, A. and Pal, C. (2007). “Improving author coreference by resource-bounded information gathering from the Web”. *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, Hyderabad, India, 2007-01-06/12, 2007 429–434.
- Kang, I.-S., Na, S.-H., Lee, S., Jung, H., Kim, P. Sung, W.-K. et al. (2009). On co-authorship for author disambiguation. *Information Processing and Management*, 45(1), 84–97.
- Katz, J. S. and Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, 40(3), 541–554.
- Ke, Q., Ferrara, E., Radicchi, F. and Flammini, A. (2015). Defining and identifying Sleeping Beauties in science. *Proceedings of the National Academy of Sciences of the United States of America*, 112(24), 7426–7431.

- Lachance, C. and Lariviere, V. (2014). On the citation lifecycle of papers with delayed recognition. *Journal of Informetrics*, 8(4), 863–872.
- Lariviere, V. and Gingras, Y. (2010a). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, 61 (1), 126–131.
- Lariviere, V. and Gingras, Y. (2010b). The impact factor's Matthew Effect: A natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, 61 (2), 424–427.
- Leimu, R. and Koricheva, J. (2005). What determines the citation frequency of ecological papers? *Trends in Ecology and Evolution*, 20(1), 28–32.
- Levitt, J.M. and Thelwall, M. (2008). Patterns of annual citations of highly cited articles and the prediction of their citation ranking: A comparison across subject categories. *Scientometrics*, 77(1), 41–60.
- Levitt, J. M. and Thelwall, M. (2009a). Atypical citation patterns in the twenty most highly cited documents in library and information science. *Scientometrics*, 78(1), 45–67.
- Levitt, J. M. and Thelwall, M. (2009b). Citation levels and collaboration within library and information science. *Journal of the American Society for Information Science and Technology*, 60 (3), 434–442.
- Leydesdorff, L. (2008). Caveats for the use of citation indicators in research and journal evaluations. *Journal of the American Society for Information Science and Technology*, 59(2), 278–287.
- Li, J. (2014). Citation curves of "all-elements-sleeping-beauties": "flash in the pan" first and then "delayed recognition". *Scientometrics*, 100(2), 595–601.
- Li, J. and Shi, D. (2016). Sleeping beauties in genius work: When were they awakened? *Journal of the Association for Information Science and Technology*, 67(2), 432–440.
- Li, J. and Ye, F. Y. (2012). The phenomenon of all-elements-sleeping-beauties in scientific literature. *Scientometrics*, 92(3), 795–799.
- Li, J., Shi, D., Zhao, S. X. and Ye, F. Y. (2014). A study of the "heartbeat spectra" for "sleeping beauties". *Journal of Informetrics*, 8(3), 493–502.
- Lindsey, D. (1989). Using citation counts as a measure of quality in science: Measuring what's measurable rather than what's valid. *Scientometrics*, 15(3/4), 189–203.
- Line, M. B. (1984). Citation decay of scientific papers: Variation according to citations received, *Journal of Information Science*, 9(2), 90–91.
- Lokker, C., McKibbin, K. A., McKinlay, R. J., Wilczynski, N. L. and Haynes, R. B. (2008). Prediction of citation counts for clinical articles at two years using data available



- within three weeks of publication: retrospective cohort study. *BMJ*, 336(7645), 655-657.
- Long, J. S. (1997). "Chap. 4 Hypothesis testing and goodness of fit". Regression models for categorical and limited dependent variables. Thousand Oaks, SAGE Publications, Inc., p.85-113.
- Lortie, C. J., Aarssen, L. W., Budden, A. E. and Leimu, R. (2013). Do citations and impact factors relate to the real numbers in publications? A case study of citation rates, impact, and effect sizes in ecology and evolutionary biology. *Scientometrics*, 94(2), 675-682.
- MacRoberts, M. H. and MacRoberts, B. R. (1987). Testing the Ortega hypothesis: Facts and artifacts. *Scientometrics*, 12(5/6), 293-295.
- MacRoberts, M. H. and MacRoberts, B. R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science*, 40(5), 342-349.
- MacRoberts, M. H. and MacRoberts, B. R. (1996). Problems of citation analysis. *Scientometrics*, 36(3), 435-444.
- MacRoberts, M. H. and MacRoberts, B. R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, 61(1), 1-13.
- Marx, W. and Bornmann, L. (2015). On the causes of subject-specific citation rates in Web of Science. *Scientometrics*, 102(2), 1823-1827.
- McCallum, A. and Wellner, B. (2003). "Object consolidation by graph partitioning with a conditionally-trained distance metric". Paper presented at the KDD Workshop on Data Cleaning, Record Linkage and Object Consolidation, Washington, DC, USA, 2003-08-24/27, 6p.
- Moed, H. F. (1989). Bibliometric measurement of research performance and Price's theory of differences among the sciences. *Scientometrics*, 15(5/6), 473-483.
- Moed, H. F. (2005). Citation analysis in research evaluation. Dordrecht, the Netherlands, Springer, 346p.
- Mryglod, O., Kenna, R., Holovatch, Yu. and Berche, B. (2013a). Absolute and specific measures of research group excellence. *Scientometrics*, 95(1), 115-127.
- Mryglod, O., Kenna, R., Holovatch, Yu. and Berche, B. (2013b). Comparison of a citation-based indicator and peer review for absolute and specific measures of research-group excellence. *Scientometrics*, 97(3), 767-777.
- 小野寺夏生, 伊神正貴. (2016). 研究計量に関するライデン声明について. *STI Horizon*, 2(4), Part.3, DOI: <http://doi.org/10.15108/stih>

- Oppenheim, C. (1997). The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British research in genetics, anatomy and archaeology. *Journal of Documentation*, 53(5), 477–487.
- Pasterkamp, G., Rotmans, J. I., de Kleijn, D. V. P. and Borst, C. (2007). Citation frequency: A biased measure of research impact significantly influenced by the geographical origin of research articles. *Scientometrics*, 70 (1), 153–165.
- Patterson, M. S. and Harris, S. (2009). The relationship between reviewers' quality-scores and number of citations for papers published in the journal *Physics in Medicine and Biology* from 2003-2005. *Scientometrics*, 80 (2), 343–349.
- Peclin, S., Juznic, P., Blagus, R., Sajko, M. C. and Stare, J. (2012). Effects of international collaboration and status of journal on impact of papers. *Scientometrics*, 93(3), 937–948.
- Peng, T.-Q. and Zhu, J. J. H. (2012). Where you publish matters most: A multilevel analysis of factors affecting citations of internet studies. *Journal of the American Society for Information Science and Technology*, 63 (9), 1789–1803.
- Persson, O., Glänzel, W. and Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432.
- Peters, H. P. F. and van Raan, A. F. J. (1994). On determinants of citation scores: A case study in chemical engineering. *Journal of the American Society for Information Science*, 45(1), 39–49.
- Rigby, J. (2013). Looking for the impact of peer review: does count of funding acknowledgements really predict research impact? *Scientometrics*, 94(1), 57–73.
- Rinia, E. J., van Leeuwen, Th. N., van Vuren, H. G. and van Raan, A. F. J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria. Evaluation of condensed matter physics in the Netherlands. *Research Policy*, 27(1), 95–107.
- Ruiz-Castillo, J. and Waltman, L. (2015). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. *Journal of Informetrics*, 9(1), 102–117.
- Schubert, A. and Glänzel, W. (2006). Cross-national preference in co-authorship, references and citations *Scientometrics*, 69 (2), 409–428.
- Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628–638.
- Seglen, P. O. (1994). Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science*, 45(1), 1–11.

- Sin, S.-C. J. (2011). International coauthorship and citation impact: A bibliometric study of six LIS journals, 1980–2008. *Journal of the American Society for Information Science and Technology*, 62(9), 1770–1783.
- Slyder, J. B., Stein, B. R., Sams, B. S., Walker, D. M., Beale, B. J., Feldhaus, J. J. and Copenheaver, C. A. (2011). Citation pattern and lifespan: a comparison of discipline, institution, and individual. *Scientometrics*, 89(3), 955–966.
- Smalheiser, N. R. and Torvik, V. I. (2009). Author name disambiguation. *Annual Review of Information Science and Technology*, 43, 287–313.
- Snizek, W. E., Oehler, K. and Mullins, N. C. (1991). Textual and non-textual characteristics of scientific papers. Neglected science indicators. *Scientometrics*, 20(1), 25–35.
- Song Y., Huang, J. Councill, I. G., Li, J. and Giles, C. L. (2007). “Efficient topic-based unsupervised name disambiguation”. Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Vancouver, Canada, 2007-06-18/23, 342–351.
- Sooryamoorthy, R. (2009). Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. *Scientometrics*, 81(1), 177–193.
- Stewart, J. A. (1983). Achievement and ascriptive processes in the recognition of scientific articles. *Social Forces*, 62(1), 166–189.
- Sun, J., Min, C. and Li, J. (2016). A vector for measuring obsolescence of scientific articles. *Scientometrics*, 107(2), 745–757.
- Tang, L. and Walsh, J. P. (2010). Bibliometric fingerprints: name disambiguation based on approximate structure equivalence of cognitive maps. *Scientometrics*, 84(3), 763–784.
- Torvik, V. I. and Smalheiser, N. R. (2009). Author name disambiguation in MEDLINE. *ACM Transactions on Knowledge Discovery from Data*, 3(3), 11:1–11:29.
- Torvik, V. I., Weeber, M., Swanson, D. R. and Smalheiser, N. R. (2005). A probabilistic similarity metric for Medline records: A model for author name disambiguation. *Journal of the American Society for Information Science and Technology*, 56(2), 140–158.
- van Dalen, H. P. and Henkens, K. (2001). What makes a scientific article influential? The case of demographers. *Scientometrics*, 50 (3), 455–482.
- van Dalen, H. P. and Henkens, K. (2005). Signals in science - On the importance of signaling in gaining attention in science. *Scientometrics*, 64(2), 209–233.
- van Raan, A. F. J. (1998). The influence of international collaboration on the impact of research results. Some simple mathematical considerations concerning the role of

- self-citations. *Scientometrics*, 42(2), 423–428.
- van Raan, A. F. J. (2004). Sleeping beauties in science. *Scientometrics*, 59(3), 467–472.
- van Raan, A. F. J. (2006). Performance-related differences of bibliometric statistical properties of research groups: Cumulative advantage and hierarchically layered networks. *Journal of the American Society for Information Science and Technology*, 57(14), 1919–1935.
- Walters, G. D. (2006). Predicting subsequent citations to articles published in twelve crime-psychology journals: Author impact versus journal impact. *Scientometrics*, 69(3), 499–510.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, vol. 10, no. 2, p. 365–391.
- Waltman, L. and van Eck, N. J. (2013a). Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison. *Scientometrics*, 96(3), 699–716.
- Waltman, L. and van Eck, N. J. (2013b). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics*, 7(4), 833–849.
- Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, 94(3), 851–872.
- Wang, J., Thijs, B. and Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PLOS ONE*, 10(5), e0127298.
- Wooding, S., Wilcox-Jay, K., Lewison, G. and Grant, J. (2006). Co-author inclusion: A novel recursive algorithmic method for dealing with homonyms in bibliometric analysis. *Scientometrics*, 66(1), 11–21.

## 謝 辞

本論文をまとめるに当たっては、多くの方のお世話になり、ここに謝意を表します。

まず、学位申請の紹介教員になってくださった影浦峽東京大学教授には、論文原稿の下書きを書いたからの突然のお願いであったにも拘わらず、快く承諾をいただき、お忙しい中、論文の構成から形式に至るまで、懇切なご指導をいただきました。学位論文の形に仕上げることができましたのは、先生のご指導によるものと感謝申し上げます。

本論文の、第3章第3節及び第4章に述べた研究の核になっている論文(1.3を参照)は共著論文であり、それらの共著者の方々に謝意を表します。第3章第3節に関する共同研究者(岩澤まり子筑波大学名誉教授、緑川信之筑波大学教授、芳鐘冬樹筑波大学教授、及び当時の筑波大学大学院図書館情報メディア研究科の6名の大学院生)の方々には、この研究で困難であった検索論文著者の識別に尽力いただいたほか、研究遂行に様々の協力をいただきました。芳鐘教授は、第4章に関する論文の共著者でもあり、重回帰分析の進め方等について大きな力をいただきました。

第3章第3節及び第4章に述べた研究は、日本学術振興会(JSPS)による科学研究費の補助(KAKENHI 助成番号 18300080 ; 平成 18~20 年度)を得て行ったものです。第5章に述べた研究は、この科研費による研究のテーマではありませんが、そこで得たデータの一部を利用しています。

最後に、家事を顧みず自分の興味あることにのみ注力する私を理解してくれた、妻の小野寺千津子に感謝したいと思います。