

令和 5 年 6 月 6 日現在

機関番号：17104

研究種目：若手研究

研究期間：2018～2022

課題番号：18K18333

研究課題名（和文）分野横断的な論文生産性指標の高精度化のための手法開発：国内研究者を対象に

研究課題名（英文）Development of a method for enhancing credibility of a normalized index to quantify publication productivity across disciplinary boundaries: a study with domestic researchers

研究代表者

山本 鉦 (Yamamoto, Koh)

九州工業大学・先端研究・社会連携本部・准教授（専門職）

研究者番号：00608903

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：正規化論文数の信頼性を高めるために、国内32大学の研究者の論文情報を収集・分析し、正規化の基準値（分野毎の平均論文生産性）を求めるうえで必要なデータ数と適切なサンプリング方法を検討した。まず偶然誤差を最小にして精度を確保するために、各分野で800以上のサンプルサイズが必要であることを明らかにした。また研究者の無作為抽出ではなく大学単位のクラスター抽出であっても、平均論文生産性の分野間の相対的な大小関係を正しく推定できることを確認し、順位などの相対的な評価に大きな影響を及ぼさないことを示した。正規化論文数の信頼性を確保するうえでは、必要サンプルサイズを満たすことが最も重要であることが示唆された。

研究成果の学術的意義や社会的意義

公開されている教員情報や論文情報を基に、正規化の基準値となる分野毎の平均論文生産性を精度よく推定する手法を示したこと、国内の研究者に関しての同様の評価を誰でも行えるよう正規化の基準値を算定し、それが分野によって系統的に異なることを数値の上から明らかにしたことには、新規性があり、学術的な意義がある。またこの取り組みは、大学などの研究機関に、さまざまな分野にまたがる所属研究者やそれらに構成される自組織を客観的、効率的に評価するための、有用な指標を提供することにつながるため、社会的な意義がある。

研究成果の概要（英文）：To enhance the credibility of a newly developed normalized index of publication productivity, publication data of researchers from 32 domestic universities were collected and analyzed; data amount and an adequate sampling method are investigated to obtain reliable standard values for normalization (disciplinary averages of publication productivity). The results showed that at least 800 sample size is required in each discipline to make random errors small and ensure precision. The sampling method applied in this study was cluster sampling at the university level, not random one at the individual level; similarity of the relative relationship of the disciplinary averages was observed in comparison with the results from a simulation of the latter method, indicating relative evaluations such as researcher rankings unsusceptible to the difference of sampling method. It suggests that meeting the criteria for sampling size is most important to ensure the credibility of the normalized index.

研究分野：計量書誌学

キーワード：正規化論文数 分野間補正 平均論文生産性 必要サンプルサイズ 層化抽出

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

昨今、大学を始めとする研究機関は国際的な競争環境下に置かれており、教育力や研究力の向上を目指したさまざまな取り組みがなされている。研究力を向上させるためには、自機関の研究力を相対的に評価することが不可欠となる。そこで、研究の活動度を表す指標として論文数が、研究の注目度や研究の質を近似する指標として被引用数が用いられている。しかし、研究分野によって論文の発表頻度や引用に関する慣習が異なり、特に被引用数は出版からの経過年にも依存するため、これらの指標の比較には注意が必要となる。分野や経過年の違いを超えて比較することを目的に、被引用数については正規化指標が開発されているものの、論文数についてはその限りではない。しかし、研究パフォーマンスの向上を目指すうえで、研究の活動度もしっかりと評価していく必要があり、そのためには論文数についての正規化指標も必要であると言える。そこで、全ての研究者が論文を執筆するうえで同等の労力（1 マンパワー）を持つという仮説の下に、各研究分野で1 マンパワーあたりの平均論文数（平均論文生産性と呼称）を求め、個人の分野別論文数をこの平均値で正規化し、分野和を求めることで得られる正規化論文数指標を開発した（Yamamoto and Ishikawa 2017）。このマンパワーという概念の導入によって、効率を考慮した正規化論文数の構築が可能となった。

2. 研究の目的

統計的に意味のある正規化論文数を構築するためには、正規化の際に必要な分野毎の平均論文生産性を、その信頼性も含めて正しく評価することが求められる。そのために必要なサンプルサイズと、適したサンプリング方法を明らかにすることが、本研究の目的である。この手法を開発することができれば、少ない労力で正規化論文数の信頼性を確保することが可能となる。

3. 研究の方法

(1) 偶然誤差を最小化するための必要サンプルサイズの検討

平均論文生産性の精度を確保するうえでは、サンプルサイズが有限であることによって生じる偶然誤差を最小化することが必要となる。ここでのサンプルとは、分析対象とする分野において各研究者が費やしたマンパワーと、そこで生産した論文数の2変数から構成される。本研究では学術文献データベースとしてScopusを活用しており、論文数はそれを基に把握し、分野分類もこれに準拠する（分野の正式名称はYamamoto and Yasunaga 2022を参照のこと）。研究者（正規的教育職員）単位で重複がないように整理したScopusの著者IDリストを有する大学に協力を依頼したり、その著者IDを公開している大学の情報を活用したりして、国内32大学に所属する約2.3万人の研究者の著者IDを収集した。これを基に個人単位で論文情報を収集し、独自のデータベースを構築した。対象期間は2011年～2020年の10年間、対象論文区分はArticleとReviewとした結果、論文数は34万報を超え、研究代表者の所属する大学と比べると80倍を超える規模となった。このデータベースを基に、各研究分野でサンプルサイズを20から、20刻みで最大サンプルサイズになるまで増やしながら、各ステップにおいてブートストラップ法で平均論文生産性の信頼区間を求め、その変化率が初めて1%以下に収束した際のサンプルサイズを、偶然誤差を最小化するための必要サンプルサイズとした。

(2) 複数大学データから得られた平均論文生産性の代表性についての検証

大学単位でクラスター抽出したデータに基づく平均論文生産性が、国内平均のよい推定値になり得るとは限らない。そこで、日本全体の論文数に対する貢献度から、日本の主要大学を4グループに分類した先行研究（村上・伊神 2018）を基に、32大学をこの4つのグループに分け、国内の平均論文生産性の推定を試みた。まず、先行研究で対象となった180を超える大学それぞれに所属する正規教育職員の数を調べ、各グループのおおよその人数シェアを求めた。その人数シェアは、貢献度が高い順に、グループ1が11%、グループ2が19%、グループ3が20%、グループ4が50%であった。次に、本研究で構築したデータベースからこの人数シェアに合うよう層化抽出したデータを用いて、国内の平均論文生産性を推定した。なお、32大学データに基づく各グループの人数は、グループ4が約4000人と最も少なかったため、これをすべて含めることとし、グループ1からグループ3の層化抽出人数は、このグループ4の人数と各グループの人数シェアを基に計算した。抽出後の研究者の総数は、8,000人程度であり、32大学全体の約35%に相当する。

4. 研究成果

(1) 偶然誤差を最小化するための必要サンプルサイズ

変化率が1%に収束するサンプルサイズは、信頼区間の上限値で800、下限値で600であった。

そのため、32大学のデータを基に、平均論文生産性を精度よく推定するためには、各研究分野において最低でも800のサンプルが必要であることが示唆された。この手法は、今後世界のデータを対象に類似の分析を行う際にも応用することができる、汎用的な手法であると考えられる。

(2) 平均論文生産性の代表性

大学単位でのクラスター抽出データに基づく平均論文生産性が、国内の平均論文生産性のよい推定値になるかを調べるため、32大学の全データに基づく値と層化抽出法による推定値を比較した(図1)。

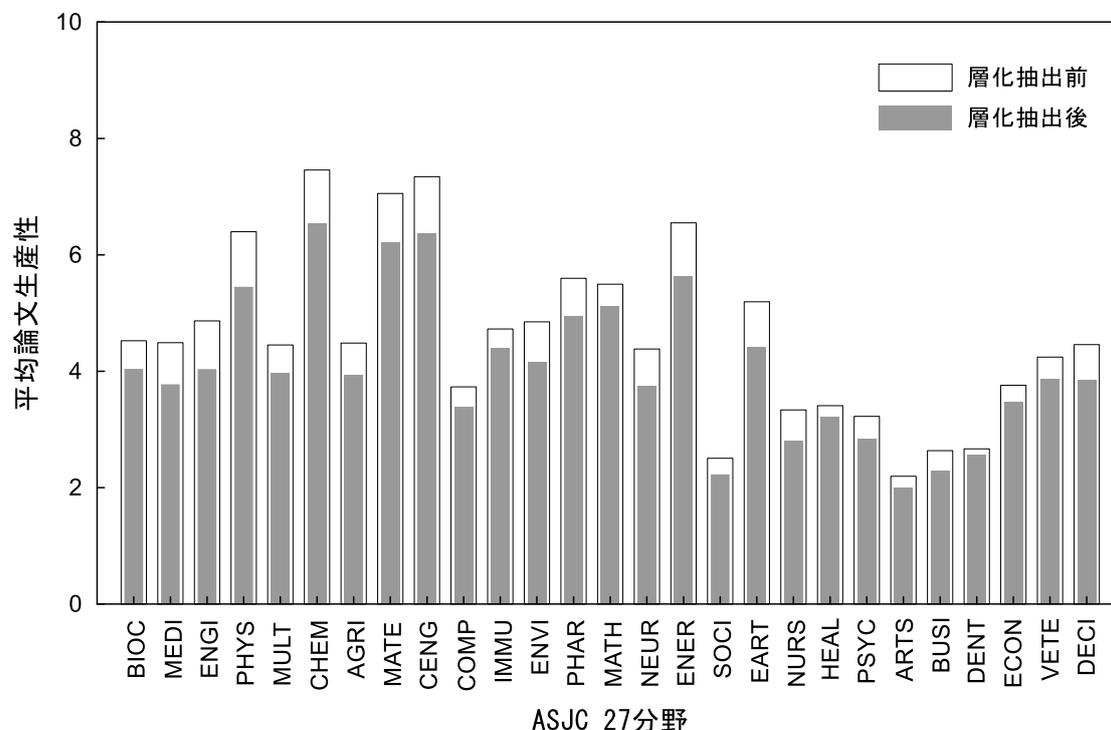


図1 各分野の平均論文生産性 (サンプルサイズ順)

層化抽出人数が全体の約35%であるため、各分野のサンプルサイズはそれぞれ最大サンプルサイズの35%程度になると考えることが妥当である。前述のように、必要サンプルサイズは800と見積もられたため、ここでは、各分野の最大サンプルサイズの35%量が800よりも大きいBIOCからEARTまでの18分野を対象に分析した。18分野の平均論文生産性の平均値は、層化抽出法を用いた場合は4.6であり、全データを用いた場合の5.2と比して約0.87倍であったが、分野間の相対的な大小関係に大きな差異は見られなかった。このことは両者の1対1プロットで、より明確に見て取ることができる(図2)。この関係が変わらなければ、正規化指標の大小にも変化はないはずである。

人数シェアが異なるにもかかわらず、平均論文生産性の分野依存性がなぜ変わらなかったのかを調べるため、各グループにおいて、そこに属するすべての研究者のデータを用いて、平均論文生産性を計算した。なお、ここでもサンプルサイズが大きい前述の18分野を分析の対象とする。4つのグループを比較すると、グループ4からグループ1にかけて、その順に全体がより高い値へとシフトしていることが分かる(図3)。下位のグループ3とグループ4の間のシフトは不明瞭であるが、グループ1からグループ3の間には明瞭なシフトが存在する。また、回帰直線の傾きもグループ間で大きな差がないことから、平均論文生産性の分野依存性は類似していることが分かる。これにより、層化抽出によって平均論文生産性の分野依存性が変わらなかったのは、論文生産性の分野間の大小関係が、大学の研究レベルによらず一定の傾向を持っていたためであると考えられる。

このことから、データセットに含まれる大学の研究レベルの偏りは、平均論文生産性のシフトを通して正規化論文数の絶対値に影響を与えるものの、それに基づく相対的な評価、例えば研究者の論文生産性順位の決定には大きな影響を与えないことが分かった。そのため、論文生産性の比較を相対的に行ううえでは、母集団全体のデータを収集したり、偏りを補正するための層化抽出を行ったりする必要はなく、平均論文生産性の精度を確保することが最も重要であり、ひいては各分野において必要サンプルサイズを満たすことが肝要であることが明らかとなった。

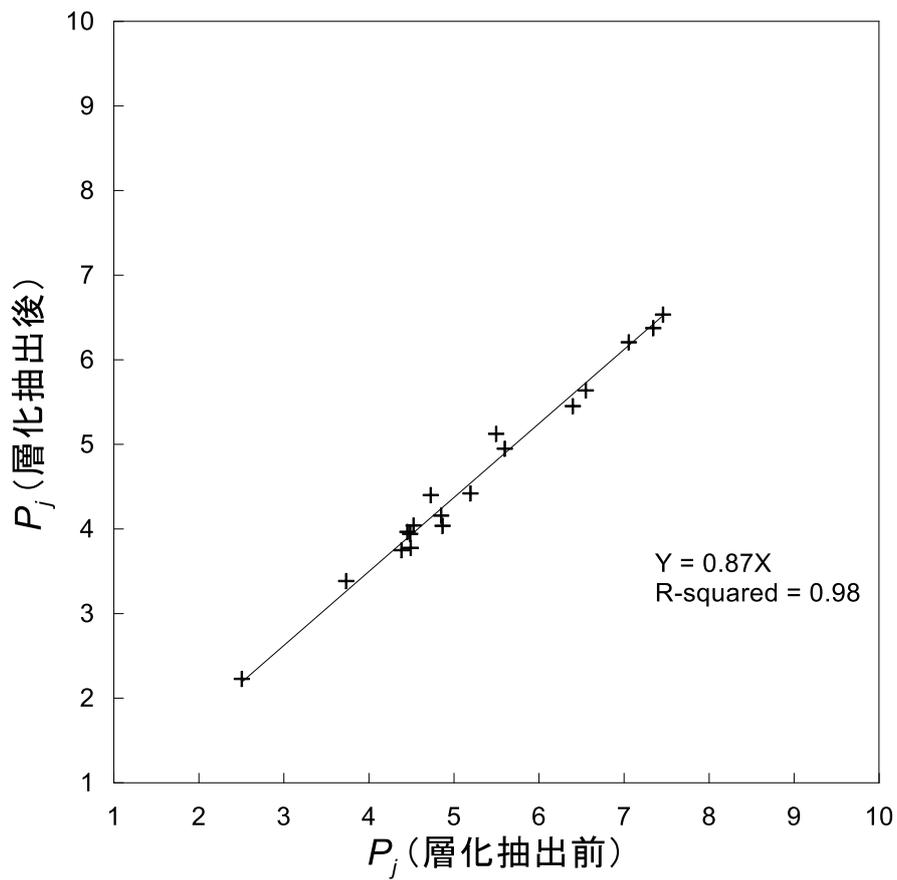


図2 層化抽出が平均論文生産性に与える影響

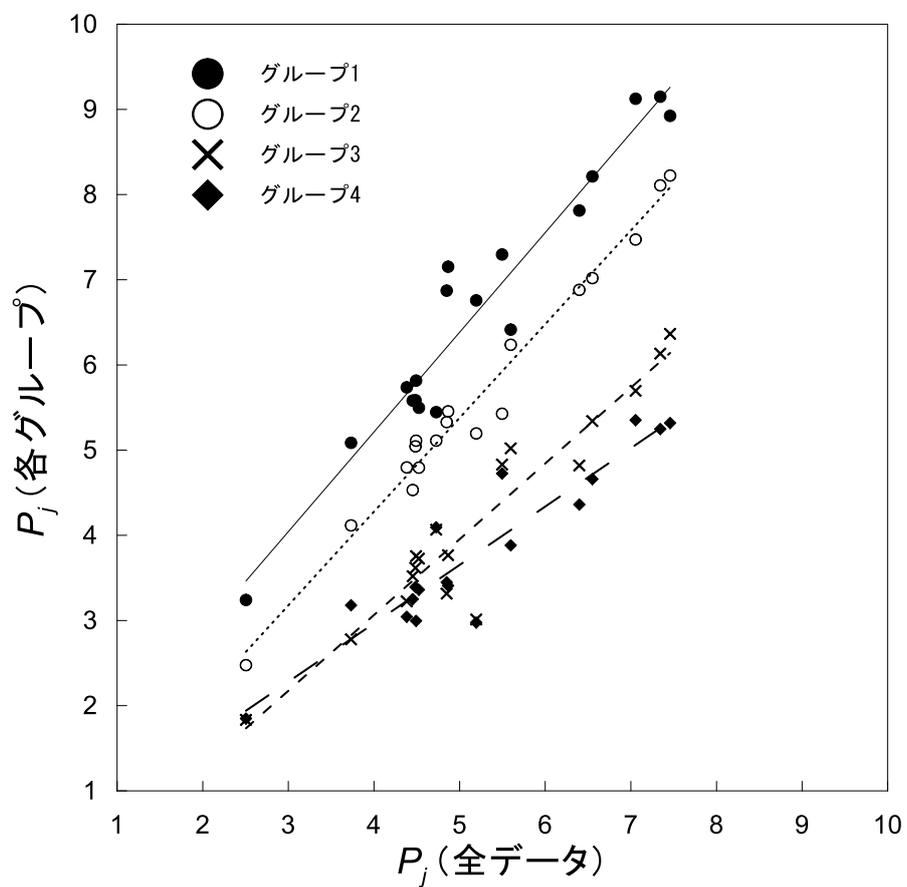


図3 4グループにおける各分野の平均論文生産性と、全データに基づく平均論文生産性

(参考文献)

K. Yamamoto and M. Ishikawa, "A normalized index to quantify publication productivity across disciplinary boundaries," *International Journal of Institutional Research and Management*, vol. 1, no. 1, pp. 67-81, 2017.

村上昭義・伊神正貫 (2018) 「日本の大学システムのアウトプット構造：論文数シェアに基づく大学グループ別の論文産出の詳細分析」, Technical report, 科学技術・学術政策研究所.

K. Yamamoto and T. Yasunaga, "A percentile rank score of group productivity: an evaluation of publication productivity for researchers from various fields," *Scientometrics*, vol. 127, no. 4, pp. 1737-1754, 2022.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Yamamoto Koh, Yasunaga Takuo	4. 巻 127
2. 論文標題 A percentile rank score of group productivity: an evaluation of publication productivity for researchers from various fields	5. 発行年 2022年
3. 雑誌名 Scientometrics	6. 最初と最後の頁 1737 ~ 1754
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s11192-022-04278-w	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 Koh Yamamoto, Takashi Okuno, Keisuke Hanabe, Teruhisa Ohno, Eitaku Nobuyama
2. 発表標題 Introduction of a normalized index for paper count into evaluation systems
3. 学会等名 Association for Institutional Research 2019（国際学会）
4. 発表年 2019年

1. 発表者名 Koh Yamamoto, Takuo Yasunaga
2. 発表標題 Introduction of a Normalized Index to Quantify Publication Productivity
3. 学会等名 7th International Symposium on Applied Engineering and Sciences（国際学会）
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

6. 研究組織

氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------