

Title	公共図書館における貸出関数の可用性の再検証
Sub Title	A re-examination of the availability of the function predicting book circulation in public libraries
Author	佐藤, 翔(Sato, Sho) 一橋, ひなた( Ichihashi, Hinata) 岡野, 真依( Okano, Mai) 川地, 美稲( Kawachi, Miina) 上嶋, 由莉( Kamijima, Yuri) 原田, 隆史( Harada, Takashi)
Publisher	三田図書館・情報学会
Publication year	2021
Jtitle	Library and information science No.86 (2021. ) ,p.19- 41
JaLC DOI	
Abstract	<p>【目的】貸出回数を被説明変数, それに影響を与えうる要因を説明変数とする「貸出関数」は, かつて日本の公共図書館を対象とする複数の研究が行われたものの, 近年では必ずしも盛んには研究されていない。本研究の目的は公共図書館の貸出関数の中でも, 自治体を単位とする線形回帰モデルによる貸出関数に着目し, その可用性を再検証することである。</p> <p>【方法】本研究では主な先行研究同様, 被説明変数として貸出密度(人口あたりの年間貸出冊数)を採用し, 説明変数の候補としては蔵書数等の図書館内の要因と, 人口密度等の自治体自体の状況を投入した, 重回帰モデルを構築する。対象とする自治体は(1)先行研究で対象とされた地域(首都圏の市等), (2)日本全国の市区町村, (3)日本の市区を8つの地域に区分した場合, とする。(1)は先行研究が対象とした地域で現在でも高い説明力のモデルが構築できるかを検証すること, (2)・(3)は先行研究で対象としていない地域でも高い説明力のモデルが構築できるかを検証することを目的とする。</p> <p>【結果】先行研究で対象とされていた地域の多くにおいては, 自由度調整済み決定係数0.7~0.8と, 先行研究に匹敵する説明力のモデルが構築できた。一方, 日本全国を対象とした場合には説明力は大きく下がり, さらに先行研究対象地域以外では, 地域を分けても必ずしも説明力の高いモデルが構築できるわけではなかった。</p> <p>Purpose: Although several studies have been conducted in the past on the function predicting book circulation in Japanese public libraries, the issue has not been studied actively in recent years. The purpose of this study is to re-examine the suitability of a linear regression model to predict the number of book circulations in public libraries per community.</p> <p>Methods: A similar model to that used in the previous studies was adopted. The dependent variable is book circulation per capita in each community. Independent variables can be divided into two groups of factors: factors related to the library, including the volume of the library collection; and community factors, including population density. There are three groups of target communities: (1) the communities targeted in previous studies, (2) all communities in Japan, and (3) cities divided into eight regions. The objective of analysis of target (1) is to verify whether it is still possible to develop a highly accurate model in the regions covered by the previous studies. The objectives in targets (2) and (3) are to verify whether we can develop the same in other regions.</p> <p>Results: It was possible to construct highly accurate models (adjusted R<sup>2</sup> are 0.7 to 0.8) in the majority of regions covered in previous studies. However, it proved</p>

	extremely difficult to construct the same when all communities in Japan were targeted. Furthermore, outside of the regions covered by the previous studies, it was not always possible to construct a highly accurate model simply by dividing communities into regions.
Notes	原著論文
Genre	Journal Article
URL	<a href="https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00003152-00000086-0019">https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00003152-00000086-0019</a>

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

## 原著論文

### 公共図書館における貸出関数の可用性の再検証

#### A Re-examination of the Availability of the Function Predicting Book Circulation in Public Libraries

佐藤 翔, 一橋ひなた, 岡野真依, 川地美穂, 上嶋由莉, 原田隆史

*Sho SATO, Hinata ICHIHASHI, Mai OKANO,  
Miina KAWACHI, Yuri KAMIJIMA, Takashi HARADA*

#### Résumé

**Purpose:** Although several studies have been conducted in the past on the function predicting book circulation in Japanese public libraries, the issue has not been studied actively in recent years. The purpose of this study is to re-examine the suitability of a linear regression model to predict the num-

---

佐藤 翔: 同志社大学免許資格課程センター, 京都府京都市上京区今出川通烏丸東入  
Sho SATO: Center for License and Qualification, Doshisha University, Imadegawa-dori Karasuma  
Higashi-iru, Kamigyo-ku, Kyoto-shi, Kyoto-fu

e-mail: min2fly@slis.doshisha.ac.jp  <http://orcid.org/0000-0001-8599-5374>

一橋ひなた: 同志社大学文学部

Hinata ICHIHASHI: Faculty of Letters, Doshisha University

e-mail: cgcc0071@mail2.doshisha.ac.jp

岡野真依: 同志社大学文学部

Mai OKANO: Faculty of Letters, Doshisha University

e-mail: cgcc5099@mail2.doshisha.ac.jp

川地美穂: 同志社大学文学部

Miina KAWACHI: Faculty of Letters, Doshisha University

e-mail: cgcc0098@mail2.doshisha.ac.jp

上嶋由莉: 同志社大学文学部

Yuri KAMIJIMA: Faculty of Letters, Doshisha University

e-mail: cgcc5041@mail2.doshisha.ac.jp

原田隆史: 同志社大学免許資格課程センター

Takashi HARADA: Center for License and Qualification, Doshisha University,

e-mail: ushi@slis.doshisha.ac.jp

受付日: 2021 年 4 月 22 日 改訂稿受付日: 2021 年 9 月 30 日 受理日: 2021 年 10 月 12 日

発行日: 2021 年 12 月 30 日

DOI: 10.46895/lis.86.19

ber of book circulations in public libraries per community.

**Methods:** A similar model to that used in the previous studies was adopted. The dependent variable is book circulation per capita in each community. Independent variables can be divided into two groups of factors: factors related to the library, including the volume of the library collection; and community factors, including population density. There are three groups of target communities: (1) the communities targeted in previous studies, (2) all communities in Japan, and (3) cities divided into eight regions. The objective of analysis of target (1) is to verify whether it is still possible to develop a highly accurate model in the regions covered by the previous studies. The objectives in targets (2) and (3) are to verify whether we can develop the same in other regions.

**Results:** It was possible to construct highly accurate models (adjusted R<sup>2</sup> are 0.7 to 0.8) in the majority of regions covered in previous studies. However, it proved extremely difficult to construct the same when all communities in Japan were targeted. Furthermore, outside of the regions covered by the previous studies, it was not always possible to construct a highly accurate model simply by dividing communities into regions.

- I. はじめに
- II. 関連研究
- III. 方法
  - A. 採用する被説明変数・説明変数とデータの出典
  - B. 分析手法
- IV. 結果
  - A. 先行研究で対象とした地域の再検証
  - B. 全国の自治体を対象としたモデルの構築
  - C. 地域を限定したモデルの構築
- V. 考察
  - A. RQ1: 先行研究対象地域における線形回帰モデルによる貸出関数の可用性
  - B. RQ2: 先行研究で対象とされていない地域、あるいは地域を限定しない線形回帰モデルによる貸出関数の可用性
  - C. 図書館・自治体の要因を用いた貸出関数の可用性と、それを高めるために今後必要な検証:「要因 X」の特定

## I. はじめに

公共図書館において、貸出回数（貸出延べ冊数）を被説明変数とし、それに影響を与えると考えられる要因を説明変数とする関数は「貸出関数」と呼ばれる。貸出関数は公共図書館の利用に関する因果関係の解明という理論面、「予測」により図書館政策・経営に有用な情報を提供しようという実践面の両面で意義を持つと考えられ、これまで複数の研究が行われてきた一方で<sup>1)</sup>、日本

においては近年、必ずしも盛んには研究されていない。本研究ではこの貸出関数、中でも自治体を単位とする線形回帰モデルによる貸出関数に着目し、その可用性をあらためて検証することを試みる。

貸出回数や来館者数といった直接的なサービス提供回数、いわゆるアウトプット指標を超えて、サービス提供がもたらした成果や価値といったアウトカム指標に注目すべきである、と言われはじめてから久しい。田辺（2016）によればすで

に1990年代末には海外の図書館評価においてアウトカム・インパクトへ注目する流れができていた。しかしながらサービスとアウトカムの因果関係が不明確である、という図書館サービスの特性もあり、アウトカム指標の図書館評価への導入はあくまで限定的であって<sup>2)</sup>、依然としてアウトプット指標は図書館において重要な役割を果たしている。図書館内での活動評価においても、あるいは対外的な説明の場でも、貸出回数の多寡や増減といったアウトプット指標は変わらず注目の的である。

しかしながら、実際にはそのアウトプット指標の妥当性についてさえ、必ずしも十分なことがわかって運用されているわけではない。例えば最も基本的な指標である貸出回数について言えば、サービスの充実などの図書館の努力や、資料費・職員数などの予算の掛け方によって左右されることが想定され、実際にそうした図書館内要因が大きな影響力を持つことが、後述のとおり先行研究から示されている。一方で、同じく先行研究から、自治体自体の状況（人口密度や住民の就業構造等）も貸出回数に影響することがわかっている。すなわち、図書館の要因のみならず自治体の状況も考慮しなければ、貸出の多寡で図書館同士を比べたり、年毎の増減について一喜一憂することには本来、意味がない。そのことには多くの図書館員・関係者は当然、気づいているはずであるが、では実際に自治体の状況等まで詳細に考慮したアウトプット指標の評価が行われているかといえば、実現していない。

自治体間・図書館間の置かれた状況の違いを考慮したり、あるいは同一自治体でも状況の変化を踏まえてアウトプット指標を運用したいのであれば、ある自治体や図書館の置かれた条件から、想定されるアウトプットの標準値を導出できればよい。その標準値に比べた実測値の高低であれば、議論する意味がある。あるいは、条件側のパラメータをどう変えれば（例えば予算や人員を増やすなど）、どうアウトプットが変化するかまで示すことができれば、図書館経営にも、政策立案にもより有効なツールとなりえるであろう。そして

このような標準値を、貸出回数を対象に導出する方法こそが、貸出関数である。

日本の公共図書館における貸出関数の検討は、後述するとおりかつて複数の研究で行われ、十分に高い決定係数を持つ（予測精度が高い）モデルを構築するに至っている。そのため一定の結論を得たと考えられたのか、近年、日本ではあまり貸出関数の研究は行われていない。しかしかつての先行研究では対象とする地域・自治体の数が限られており、また研究実施から約30年が経過しているため、現在でも先行研究のモデルが当てはまるのかは未知数である。加えて、計算機能力の向上等により、データさえあれば追試は容易にもなっている。

そこで本研究では全国の自治体を対象とし、線形回帰モデルを用いた貸出関数による予測の可用性を再検証する。もし全国の自治体を対象とする、精度の高いモデルを構築することができたならば、このモデル自体や、各自治体にあてはめた検証結果（モデルによる予測と実際の差）を公開することで、日本の公共図書館評価に関する新地平を切り拓くことができるであろう。逆に精度の高いモデルを構築できないとすれば、その理由を検証することで、図書館利用（本研究の場合は貸出）に影響を与える要因に関する、新たな知見を得ることができると考えられる。

以上の背景と目的意識の下で、本研究では以下のリサーチ・クエスション（RQ）を設定し、分析を進めていく。

RQ1：先行研究で対象とされていた地域において、現在でも線形回帰モデルにより、高い決定係数を持つモデルが構築できるか

再検証にあたっては、まず先行研究が対象としていた地域において、現在でも同様に線形回帰モデルによる、貸出回数の予測が実現可能であるかを検討する必要があると考えた。これがそもそも再現できないのであれば、本研究の前提が破綻する。当時と同様に高い決定係数を持つモデルが構築できる、すなわち精度の高い貸出回数の予測が可能であるならば、次の段階である全国単位での

検証に入る意義はあると考えられる。

RQ2: 先行研究で対象とされていない地域において、線形回帰モデルにより、高い決定係数を持つモデルは構築できるか

RQ1 について、先行研究が対象としていた地域におけるモデルの再現性が確認できたのであれば、対象とする地域を広げ、他の地域や全国規模でもモデルが成り立つのかを検証する価値があると考えられる。この RQ2 はさらに 2 つの下位レベルに分けられる。

RQ2-1: 地域を限定しないモデルの構築は可能か

RQ2-1 ではまず、全国規模で貸出回数予測の高精度で可能かを検証する。後述のとおり、先行研究では対象の質・条件を揃えるために地域を限定していたが、本当に地域を限定しなければ高い決定係数を持つモデルは構築できないのか、の検証とも位置付けられる。

RQ2-2: 地域を限定したモデルの場合、地域ごとに妥当なモデルは構築できるか

RQ2-1 については、おそらく高精度での予測は実現できないことが考えられる。その場合、先行研究で対象とした地域以外でも、地域を限定すれば高い精度での貸出回数の予測は可能なのか（具体的には、先行研究は首都圏等を対象としていたが、北海道や東北、四国等では成り立つのか）、ということが次の疑問として湧く。RQ2-2 ではこの疑問について検討していく。

## II. 関連研究

貸出関数のように、図書館の利用に影響を与える要因を特定し、利用量を説明するモデルを構築する試みは、これまで数多く行われてきている。そのようなモデルについてレビューした岸田 (1986) は、調査の単位としては個人（ある個人が図書館を利用するか、あるいは利用頻度の予測を行う）と、コミュニティあるいは図書館（あるコミュニティ・図書館における利用量の予測を行う）があることを示している<sup>3)</sup>。そのほかには

資料の単位に着目する（資料ごとの利用予測を行う）場合や、同一図書館について日毎の予測を行う（ある日の利用量を予測する）場合もありうるであろうが、このうち本研究で着目する貸出関数は、コミュニティあるいは図書館単位のモデルである。

このうち図書館を単位とする予測について、初期の例としては米フロリダ州の 32 の郡図書館を対象に、登録利用者数、サービス対象者あたりコスト、サービス対象者あたり蔵書数といった図書館に関する要因と、教育レベルという自治体の状況の要因を説明変数とし、重回帰分析により貸出密度の説明を試みた Kim・Shin (1977) がある。この研究では前述の 4 つの説明変数により、決定係数 0.74 と一定のレベルで貸出密度を予測することが可能であったとしており、特に影響度が大きかったのは教育レベル、次いでサービス対象者あたり蔵書数であった<sup>4)</sup>。日本における図書館単位の予測の初期の例としては、全国 400 の市区町村立図書館を対象に、公共図書館の活動指標と、図書館の内的要因の関係を分析した糸賀 (1982) がある。糸賀の研究では公共図書館の活動指標のうち、因子分析の結果から 4 つの代表的指標（貸出便益、登録率、蔵書回転率、実質貸出密度）を特定した上で、図書館の内的要因からそれぞれの代表的指標を説明する重回帰分析を行っている。このうち貸出に関する指標の結果としては、実質貸出密度（登録者数あたりの貸出冊数）と蔵書回転率（蔵書あたりの貸出冊数）があり、実質貸出密度については貸出条件や登録者 1 人あたり購入冊数、蔵書回転率は蔵書に占める児童書の比率や蔵書新鮮度の影響が大きいことが指摘されている。また、対象とする 400 館すべてを含む重回帰分析では重相関係数が低く、蔵書数から対象を 3 グループに分ける必要があったとも述べられている<sup>5)</sup>。さらに、糸賀と同じデータを用いつつ、対象を東京 23 区の図書館に、被説明変数を貸出密度（個人貸出）に限定し、サービス対象人口や図書館同士のサービス圏域の重なり等の新たな要因も考慮した重回帰分析を実施した常盤 (1985) は、サービス対象一人当たり受入数、他館とのサー



ビス圏域の重なり、団体・移動貸出、貸出制限冊数、建設年の5つの説明変数を用いることで、決定係数0.812とある程度のレベルで貸出密度の予測が可能であることを示した<sup>6)</sup>。

より本稿に関連するコミュニティを単位とする予測に関して、最初期の例としては米国の地域（コミュニティ）を単位とし、成人男性の教育状況、女性の教育状況、年齢、住民に占める白人の割合といった地域の状況から、公共図書館における貸出密度の説明を試みたParker・Paisley（1965）がある。この研究では女性の教育状況が最も影響力の大きい説明変数であり、それに人口と収入状況のデータを加えることで、重相関係数 $R=0.38$ で貸出密度を予測できるモデルが構築でき、その他の変数を足しても説明力は大きく向上はしなかったと報告されている<sup>7)</sup>。また、海外における近年の研究例としては米マサチューセッツ州の280のコミュニティを対象に、コミュニティが図書館に割り当てる予算と貸出、図書館のそのほかのアウトプット（提供するプログラム数、来館者数、所蔵資料数、職員の総勤務時間数。所蔵資料数と職員の総勤務時間数についてはアウトプットではなくインプットとみなすべきとも考えられるが、著者はアウトプットとして挙げている）と貸出、コミュニティの社会経済状況と図書館予算との関係を分析した、Carlozzi（2018）がある。この研究では貸出と予算の間には相関係数0.93と非常に強い相関があることを指摘しており、また各アウトプットによって貸出を決定係数0.87と比較的、高い水準で説明できることを示している。中でも職員の勤務時間数の影響が最も大きく、勤務時間数は予算とも特に高く相関している（労働時間が増えれば予算もかかることを考えれば当然である）ことから、予算と貸出冊数の間に強い関係があることが示されている<sup>8)</sup>。さらに図書館予算は地域の教育状況の影響を強く受けることも示しているが、この研究では貸出冊数を直接的に、地域の社会経済状況と結びつけて分析することは試みられていない。そのほかに貸出冊数のみを対象とするものではないが、人口あたりの貸出冊数を含んだ複数の図書館活動に関す

る指標を組み合わせた、なんらかの新指標を算出し、これを地域の社会経済状況を説明変数とする重回帰分析によって説明することを試みた研究として、Seavey（1987）とKoontz（1992）がある（具体的に新指標算出に用いる個々の指標は両研究で異なるが、いずれも人口あたり貸出冊数は含んでいる）。このうちSeavey（1987）ではウィスコンシン州内の地域を対象に、1970年時点と1980年時点での新指標の説明モデルを構築し、いずれも最も影響の大きい説明変数は都市住民の割合、2番目は経済状況、3番目は人口あたり支出であった、としている。この3つの指標を用いたモデルの決定係数は1970年時点で0.525、1980年時点で0.66であった<sup>9)</sup>。また、Koontz（1992）では国の6つの地域それぞれを対象として新指標の説明モデルを構築しており、うち3つで決定係数0.8以上、2つで0.6以上と一定の説明力が示されているものの、残る1つでは0.44と説明力は限定的であった。また、人種構成、教育状況、収入状況等が新指標に強く影響する傾向はあったものの、地域ごとにモデルに含まれる説明変数には異なる傾向も存在した<sup>10)</sup>。

日本において、コミュニティを単位とする利用予測の研究に先鞭をつけたのは岸田（1986）である<sup>3)</sup>。同研究では首都圏（東京都、埼玉県、千葉県、神奈川県）の市と特別区を対象に、自治体（市・特別区）単位での貸出密度（人口あたりの年間貸出冊数。ただし児童書と児童人口を除く）を算出し、35の説明変数を用いた重回帰分析によってこれを予測するモデルを構築した。その結果、最終的に「蔵書数」、「専門職従事者数」など、10の説明変数から貸出密度を予測するモデルが構築され、決定係数0.809、自由度調整済み決定係数0.791と高い説明力を示した。さらに因子分析により説明変数のカテゴリー分けを試みたところ、「知的水準」、「都会化」、「地域繁栄度」、「図書館活動」、「成長中であること」の5つの因子が特定され、それぞれ比較的均等に最終的なモデルに取り込まれていた。その中でも最も影響力が強いのは「図書館活動」（蔵書数等が含まれる）であり、次が「知的水準」（専門職従事者数など

が含まれる)で、その2つの因子で変動のかなりの部分が説明できた、とされている。岸田の研究は1980年度という一時点のものであったが、春日部(1990)は東京都の市・特別区に限定し、1965・1970・1975・1980・1985年の5時点について、同様に貸出密度を被説明変数とする重回帰分析を実施している。その結果、いずれの時点についても決定係数0.8以上と高い説明力が示された。貸出密度に一貫して大きく影響する要因としては、図書館の内的要因としては蔵書数と職員1,000人当たり人口、外的要因としては昼間人口と専門職・管理職があげられている。しかし、全体に説明変数のうちいずれが高い偏相関を示すかは時点によって大きな変動があることも指摘されている<sup>11)</sup>。さらに岸田・佐藤(1991)では首都圏以外での同様のモデルの構築可能性を検証するため、大阪府と富山県の自治体を対象に、それぞれ1980年度・1985年度という2時点について、重回帰分析によるモデル構築を試みている。その結果、1985年度の富山県については自由度調整済み決定係数0.3909と十分な説明力を持つモデルが構築できなかったものの、その他は自由度調整済み決定係数0.6892~0.7931と、高い説明力を持つモデルが構築された。ただし、大阪府において高い決定係数を持つモデルが構築される一方、富山県は1980年度も大阪府に比べると決定係数の低いモデルしか構築されなかったことから、首都圏や大阪府といった大都市においてのみ、重回帰モデルが高い説明力を持つのではないかと、いう指摘もなされている。また、先行研究も含めて構築されたモデルのいずれもが、最終的に異なる説明変数を投入するモデルであったことから、利用に関する因果関係の解明としても、実用的な「予測」としてもこのままでは役に立たない、という指摘もなされている。岸田・佐藤らはこれを説明変数間の多重共線性の影響と判断しており、変数選択をデータに基づき帰納的に行うアプローチの限界であるとしている<sup>1)</sup>。そのような判断もあつてか、以降、岸田は帰納的アプローチではなく、演繹的に貸出関数のモデルを導出する方向に舵を切り、岸田(1998)では蔵書数と定住人口の

みから貸出延べ冊数を説明するモデルを構築している<sup>12)</sup>。さらに岸田(2013)では検討を進め、計量書誌学的要因(各図書館の貸出回数は広義のZipfの法則に従う)と地理的要因(住民のうち、図書館を利用している人の割合は、図書館からの距離に逆比例する)を組み込んだ貸出予測式を構築している。この予測式は少ない前提で高い決定係数を示したが、被説明変数は貸出密度ではなく貸出冊数そのものであるため、貸出密度を対象とする先行研究の回帰分析との単純な比較は困難である<sup>13)</sup>。

岸田ら以降の貸出関数に関する研究として、田村(1994)は人件費、物品費および登録者数を説明変数とし、貸出冊数を予測するモデルを構築している。岸田・佐藤らまでの先行研究が線形回帰(重回帰分析)を用いているのに対し田村は対数線形回帰を用いている点、自治体人口別に分析することで、自治体規模によって異なる傾向を示している点に特徴があり、人口の少ない自治体では人件費・物品費の変動が貸出冊数の変動要因となる(人件費・物品費を増やせば貸出冊数も増える)のに対し、人口がある程度の規模以上になると人件費が固定費的になり(人件費を増やしても貸出冊数は増えない)、さらに人口が多い自治体では人件費も物品費も固定費的になる(いずれを増やしても貸出冊数は増えない)ことを示している<sup>14)</sup>。また、田村(2004)は経済状況の変化と公共図書館の利用の関係に関する回帰分析も行っている。この研究では都道府県単位に集計したデータを用い、有効求人倍率・インフレ率と貸出密度の関係についてのパネルデータ回帰分析を行い、経済状況の悪化が公共図書館の利用を増加させることを示している<sup>15)</sup>。

近年の先行研究として、金井(2017)は全国の自治体(人口100万人未満の市と特別区)を対象に、自治体を単位とし、貸出冊数(密度ではなく冊数そのもの)を被説明変数とする、5つの線形回帰モデルを検証している。結果から、蔵書数と奉仕人口のみでも説明力の高い(自由度調整済み決定係数0.8668)モデルが構築でき、人口密度も加えればさらに説明力が上がる一方、開架冊数や



登録者数を加えることはほとんどモデルを改善しないことを示している<sup>16)</sup>。貸出密度ではなく冊数を単位としている点で、蔵書数と人口のみで十分な説明ができるのは当然ともいえるが（つまり、単純に規模が大きい自治体・規模が大きい図書館は、貸出冊数も多い）、「借りやすいかどうか」（開架冊数）や「資料を借りる人」の多寡（登録者数）がモデルに寄与しない可能性を示したことは有効である。

岸田らが指摘するように変数選択をデータに基づき、帰納的に行うアプローチには限界がある。一方で、貸出関数を演繹的に導くための、検討材料としては、サンプルの大きなデータを用いた、帰納的な推論と検証の積み重ねもまた必要であることも、岸田らは指摘している<sup>1)</sup>。それにも拘わらず以降、日本国内において、大きなサンプルデータを用いた、単純な帰納的アプローチは行われてこなかった。本稿で報告する研究結果はまさにそのような、よりサンプルを拡大した、帰納的アプローチの検証として位置づけられる。

### III. 方法

#### A. 採用する被説明変数・説明変数とデータの出典

自治体を対象とする貸出関数の検証においては、被説明変数として貸出密度（人口あたりの年間貸出冊数）を採用する場合がほとんどである。単純な貸出回数を説明変数とした場合、前述のとおり、自治体規模の影響が非常に大きくなることは当然であり、それ以外の効果が見えにくくなる。よって本研究でも貸出密度を被説明変数として採用することとした。さらに、日本における主要な3つの先行研究（岸田（1986）<sup>3)</sup>、春日部（1990）<sup>11)</sup>、岸田・佐藤（1991）<sup>1)</sup>）では児童書の貸出については成人向け資料の貸出とは傾向が異なると考え、児童書の貸出や児童人口（15歳未満人口）を除いたデータを用いて貸出密度を算出している。先行研究の有効性の再検証という本研究の目的を鑑み、本研究でも同様に児童書と児童人口を除いた貸出密度を被説明変数とした。

説明変数については、主要な先行研究におい

て図書館内の要因（蔵書数、受入数、資料費等）と、自治体自体の状況（女性の割合、教育状況、専門職従事者の割合、昼間人口等）の両方を採用し、高い説明力を持つモデルを構築している。本研究も先行研究同様、図書館内の要因と自治体自体の状況の双方に関する説明変数を投入することとする。

被説明変数である貸出密度の計算に用いる年間貸出冊数、および図書館内の要因に関する説明変数のデータは『日本の図書館 統計と名簿』2018年版（電子版）<sup>17)</sup> から、自治体自体の状況に関する説明変数データは2015年国勢調査<sup>18)</sup> から取得した。先行研究のうち岸田（1986）<sup>3)</sup> では朝日新聞社や東洋経済新報社の資料等も使い、「テレビ普及率」や「新聞頒布数」、「都市化」、「民力」等といった指標も説明変数に採用している。しかしそれらの指標の中でモデルに大きく寄与したものは必ずしも多くはない。春日部の先行研究では国勢調査以外を出典とする自治体に関するデータとして書籍・雑誌販売額を投入しているが<sup>11)</sup>、これも必ずしもモデルに寄与していない。それもあるのか、岸田・佐藤（1991）では自治体状況に関する説明変数として国勢調査のデータのみを用いるようになっており<sup>1)</sup>、地域によっては十分に高い説明力を持つモデルを構築した。国勢調査以外を出典とするデータは取得の労に対してモデルへの寄与が大きくはないと考えられ、本研究でも国勢調査のデータのみを用いることとした。さらに、国勢調査は10年ごとに行われる調査項目数の多い大規模調査と、5年ごとに行われる項目数の少ない簡易調査があり、岸田（1986）の先行研究<sup>3)</sup> では大規模調査（1980年）のデータを用いているため地域住民の学歴等のデータも説明変数として採用しているが、簡易調査の年のデータも用いている春日部<sup>11)</sup> や岸田・佐藤<sup>1)</sup> では大規模調査にしか出現しない項目は採用していない。本研究も、より直近の状況を検証することを目的としており、直近の国勢調査（2015年）は簡易調査であるため、大規模調査にしか出現しない項目は説明変数としては採用できないこととなった。以上の限界はあるものの、それ以外は先行研究で

用いられている説明変数はできるだけ採用する方針とし、最終的に第1表の21個の説明変数を採用することとした。

なお、本研究では図書館に関するデータの出典とする年（2018年）と、自治体状況のデータの出典とする年（2015年）の間に3年間のずれがある。本来であれば出典とするデータの時期はできるだけ近づけるべきであり、前述の主要な先行研究でも、ずれても1年前後の範囲内に収めている。本研究でデータの出典がずれた理由は、採

用した被説明変数にある。本研究では前述のとおり、先行研究に揃え、児童書と児童人口を除いた貸出密度を被説明変数として採用することとした。しかし『日本の図書館統計と名簿』では2017年版まで、児童書の蔵書数や、貸出冊数のうち児童書が占める割合など、児童サービスに関する項目を調査対象から除いていた。これら児童サービスに関する項目が質問票に復帰したのは2018年版からである。図書館データと自治体状況に3年間のずれがあることと、そもそも被説明

第1表 投入する説明変数候補

項番	変数名	算出方法	採用先行研究
1	女性の割合	女性の人口／全人口	すべて
2	非労働者の割合	(成人（15歳以上）人口－15歳以上就業者数）／成人（15歳以上人口）	すべて
3	専門職従事者の割合	専門職従事者数／全就業者数	すべて <sup>1)</sup>
4	管理職従事者の割合	管理職従事者数／全就業者数	すべて <sup>1)</sup>
5	第1次産業就業者割合	国勢調査データから直接取得	すべて
6	第2次産業就業者割合	国勢調査データから直接取得	岸田（1986）、春日部（1990）
7	第3次産業就業者割合	国勢調査データから直接取得	すべて
8	昼間人口	昼間人口／定住人口	すべて
9	人口密度	国勢調査データから直接取得	すべて
10	蔵書数	(全蔵書数－児童書蔵書数)／(全人口－児童人口 <sup>2)</sup> )	すべて
11	受入数 <sup>3)</sup>	(全受入数－児童書受入数)／(全人口－児童人口 <sup>2)</sup> )	すべて
12	資料費	図書館資料費／全人口	すべて
13	職員数	図書館職員数／全人口	岸田（1986）、岸田・佐藤（1991）
14	移動図書館の有無	—	岸田（1986）
15	移動図書館数	—	春日部（1990）、岸田・佐藤（1991）
16	図書館面積	—	すべて
17	図書館数 <sup>4)</sup>	—	すべて
18	図書館密度	図書館数／自治体面積	すべて
19	雑誌種数 <sup>5)</sup>	雑誌購入種数／全人口	春日部（1990）
20	開館時間 <sup>6)</sup>	中央図書館の開館時間を分単位で算出	春日部（1990）
21	職員1000人あたり人口	全人口／職員数＊1000	春日部（1990）

- 1) 春日部（1990）、岸田・佐藤（1991）は専門職従事者と管理職従事者数の合算値を投入している。
- 2) 岸田（1986）は児童を12歳未満、春日部（1990）は10歳未満としているが、本研究では国勢調査データの整形の都合上、15歳未満としている。
- 3) 春日部（1990）、岸田・佐藤（1991）は受入数について、全受入数を全人口で除しているが、本研究では蔵書数との一貫性を重視し、岸田（1986）にならって児童書・児童人口を除いて受入数を算出している。
- 4) 岸田・佐藤（1991）は図書館数を自治体人口で除しているが、他の先行研究は単純な館数である。本研究では後者を採用している。
- 5) 春日部（1990）は雑誌種数を図書館数で除し、さらに人口で除しているが、本研究では自治体単位で集計済みの購入種数を、単純に人口で除している。
- 6) 春日部（1990）は年間の延べ開館時間を月数(12)で除し、それを24(1日)で除し、さらに図書館数で除している。本研究では中央図書館に限定し、1日あたりの開館時間を開館/閉館時刻から分単位で算出した上で、開館日数を乗じている。

変数が異なってしまう（児童書・児童人口を除けなくなる）場合のいずれが先行研究の再検証に大きな影響を及ぼすかを検討した結果、データ取得時期のずれの方が影響が小さいであろう（3年間で極端に変わることはないであろう）と判断し、今回はデータ取得年のずれは受け入れることとした。2020年国勢調査の結果が公開されれば、この問題は解決されると期待したが、2020年は新型コロナウイルス感染症の影響もあって図書館側の利用状況が安定しないと考えられ、国勢調査と図書館統計の取得時期を完全に一致させ、先行研究の検証ができるのは2025年以降を待たねばならない。

## B. 分析手法

### 1. RQ1について

RQ1では先行研究で対象とされていた地域において、現在でも線形回帰モデルにより、高い決定係数を持つモデルが構築できるかを検証する。検証の対象とする先行研究は前述の主要な先行研究、岸田（1986）<sup>3)</sup>、春日部（1990）<sup>11)</sup>、岸田・佐藤（1991）<sup>1)</sup>であり、対象とする自治体はそれぞれ首都圏（東京都、埼玉県、千葉県、神奈川県）の市区、東京都の市区、大阪府と富山県の市町、である。先行研究実施当時と現在では市区町村合併等により対象となる自治体の数は変化している可能性があるが、その変化を踏まえて現状がどうなっているかを検証すべきと考え、本研究では2015年時点の、各地域における先行研究で対象とした自治体種別を、RQ1の検証対象とすることとした。

先行研究のうち岸田（1986）<sup>3)</sup>は説明変数の選択にステップワイズ法、岸田・佐藤（1991）<sup>1)</sup>は変数減少法を用い、春日部（1990）<sup>11)</sup>では強制投入法を用いている。本研究の検証でも先行研究でステップワイズ法を用いていた地域ではステップワイズ法、変数減少法を用いていた自治体では変数減少法、強制投入法を用いていた地域では強制投入法を用いる。ステップワイズ法と変数減少法を用いた場合、先行研究と同じ説明変数が選択されるとは限らない（先行研究の一部の説明変数が

とれていないことを考えれば、当然選択されない）が、どのような変数が選択されるかに依らず、「ステップワイズ法あるいは変数減少法を用いて高い決定係数を持つモデルが構築できる」という事実が再現できるか自体をここでは検証する。

### 2. RQ2について

RQ2では先行研究で対象とされていなかった地域について、線形回帰モデルにより、高い決定係数を持つモデルが構築できるかを検証する。

具体的にはまずRQ2-1として、地域を限定しない、すなわち全国の自治体を対象とするモデルの構築を試みる。この際、全自治体種別を対象とする場合、自治体種別を分けた場合（市区、町、村）のモデルをそれぞれ作成し、また自治体種別市区については先行研究にならって、人口100万人未満に限定した場合も作成する。

「I. はじめに」でも述べたとおり、RQ2-1で作成する地域を限定しないモデルは、地域を限定した先行研究に比べて決定係数は低くなることが予想される。そもそも先行研究で地域が限定されていたのは、“重回帰分析で良い結果を得るためにはある程度その対象の質ならびに条件を揃えることが必要”<sup>3)</sup>との認識に基づく。つまり説明変数には表れていない、なんらかの潜在的な質・条件があり、かつそれは地域や自治体種別を限定することで揃えることが可能、と先行研究では想定している。実際に首都圏以外、大阪府や富山県を対象とする分析でも一定のモデルが構築されていることを鑑みればこの想定は妥当と考えられるが、未だにそれら以外の地域での検証は行われていない。一方、都道府県単位での実施はあまりにモデルの数が増え、また含まれる自治体数が少なくなりすぎるため、結果の解釈が困難である。そこで本研究ではRQ2-2として、都道府県を北海道、東北、関東、中部、近畿、中国、四国、九州の8つに区分し、8地域ごとに一定の決定係数を持つモデルが構築できるかを検証することとした。対象とする自治体種別もそろえるため、ここでは市と区に限定して分析を行った。

なお変数選択についてはRQ2-1、RQ2-2ともに

ステップワイズ法を採用した。強制投入法も実施し、両手法の結果を併記することも考えたものの、それでは紙数をとりすぎるため、より多くの先行研究で採用されているステップワイズ法を用いることとした。

#### IV. 結果

##### A. 先行研究で対象とした地域の再検証

###### 1. 岸田（1986）：首都圏の市区

第2表に岸田（1986）<sup>3)</sup> 同様、首都圏（東京都、埼玉県、千葉県、神奈川県）の市区を対象に、ステップワイズ法を用いて構築した、貸出密度を説明する重回帰モデルを示す。ステップワイズのためのF値確率は投入を0.05、除去を0.10とした（以下、ステップワイズ法を用いたモデルについてはすべて同様）。対象となった自治体数は126市区であった（岸田（1986）では116市区）。なお、開館日数を0と回答している自治体については、実際に閉館しているとは考えにくいものの、確証が得られないため本研究では分析に加えていない。児童への貸出冊数を0と回答している自治体も、児童サービスに関する回答が実質、無回答であったと判断し、分析に加えていない（開館日数0、児童への貸出0の自治体の扱いは以下、同様）。

最終的には7つの説明変数で、重相関係数0.900、決定係数0.809、自由度調整済み決定係数0.798のモデルが構築された。岸田（1986）で構築されたモデルは10の説明変数を用い、重相関

係数0.899、決定係数0.809、自由度調整済み決定係数0.791であり、採用された説明変数は後述のとおり大きく異なるものの、説明力についてはほとんど変わらないモデルが構築できたといえる。

採用された説明変数のうち、もっとも影響が大きいのは職員数で、次いで受入数、昼間人口の影響が大きい（昼間人口は係数が負）。図書館に関する要因としてはそのほかに蔵書数、開館時間が、自治体の状況に関する要因としては第3次産業就業者割合、管理職従事者割合が採用された。このうち岸田（1986）でも採用されていた説明変数は受入数と蔵書数のみである。開館時間は岸田（1986）にはなかったのが当然として、自治体の状況に関する要因としても、岸田（1986）では影響が大きかった専門職従事者割合や人口密度が含まれていない。このように含まれる説明変数が大きく異なるにも拘わらず、結果的に先行研究とほとんど同じ説明力を持つモデルが構築できたことは、後の先行研究でも指摘されていたとおり、説明変数間の多重共線性の問題をうかがわせる（ただし多重共線性の目安となるVIFについては、第2表のとおり採用された変数はいずれも4未満にはとどまっている）。よって個々の説明変数として何が採用されたかを重視することにはあまり意味がない。先行研究から30年以上を経ても、首都圏の市区においては未だ、自治体側の要因と図書館側の要因を用いて、高い決定係数を持つ貸出密度の説明モデルが構築可能である、という事実のみをここでは押さえておくべきであろう。

第2表 岸田（1986）貸出密度予測再検証モデル（首都圏の市区126自治体対象・ステップワイズ法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
職員数	4.085	0.357	0.380	4.468	0.000	3.945
第3次産業就業者割合	0.073	0.199	0.286	3.237	0.002	2.337
蔵書数	0.325	0.195	0.271	3.064	0.003	2.518
受入数	11.788	0.350	0.379	4.453	0.000	3.829
昼間人口	-0.643	-0.314	-0.390	-4.595	0.000	2.881
開館時間	0.000	0.130	0.249	2.787	0.006	1.339
管理職従事者の割合	42.759	0.156	0.204	2.267	0.025	2.935
重相関係数	0.900					
決定係数	0.809					
自由度調整済み決定係数	0.798					



## 2. 春日部（1990）：東京都の市区

第3表に春日部（1990）<sup>11)</sup> 同様、東京都の市区を対象に、強制投入法を用いて構築した、貸出密度を説明する重回帰モデルを示す。対象となった自治体数は42市区であった（春日部（1990）では1965年が29、1970年が30、1975年が46、1980年が47、1985年が49。児童サービスに関する無回答自治体の存在により、本研究では含まれる自治体数が減少したと考えられる）。なお、春日部（1990）については専門職従事者と管理職従事者の割合の合算値を説明変数として採用しているため、本項のモデルでも合算値を投入している。

強制投入法を用いているが、許容量が基準を下回ったため、第3次産業就業者割合と移動図書館数はモデルから除外された。また、春日部（1990）には存在した「主婦の割合」と「書籍・雑誌販売額」が前述の制約により、本研究では含まれていない。そのためモデルに含まれる説明変数は17と春日部（1990）の20より3つ少ないが、本研究のモデルの重相関係数は0.941、決

定係数は0.886、自由度調整済み決定係数は0.805と比較的高い。5年区切りで分析を行った春日部（1990）のうち、最も決定係数の高い1960年で決定係数0.95、最も低い、かつ分析に含まれる最新の年である1985年で0.82であった。それから35年が経過しているにも拘わらず、当時と同じか少し少ない説明変数を用いて、説明力がほとんど変わらないモデルを構築可能であるといえよう。

各説明変数について、春日部（1990）では図書館に関する要因としては蔵書数と職員1000人当たり人口、自治体の状況としては昼間人口、専門職・管理職従事者の割合、そして女性の割合が一貫して高い影響力を示した、とされている。本研究の結果でも職員1000人当たり人口は比較的高い影響力を示しているが、図書館に関する要因で最も影響力が大きいのは受入数であり、他にも雑誌種数や資料費の方が蔵書数よりも高い影響力を示している。自治体に関する要因も、最も影響力が大きいのが昼間人口であること、女性

第3表 春日部（1990）貸出密度予測再検証モデル（東京都の市区42自治体対象・強制投入法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
女性の割合	62.344	0.227	0.353	1.851	0.077	3.158
非労働者の割合	7.794	0.091	0.192	0.960	0.347	1.890
専門職・管理職従事者の割合	3.696	0.057	0.063	0.308	0.761	7.152
第1次産業就業者割合	-1.459	-0.247	-0.249	-1.259	0.220	8.104
第2次産業就業者割合	0.104	0.176	0.184	0.917	0.368	7.748
第3次産業就業者割合	除去					
昼間人口	-0.819	-0.631	-0.458	-2.523	0.019	13.131
人口密度	0.000	-0.138	-0.136	-0.673	0.507	8.799
蔵書数	0.293	0.164	0.165	0.818	0.421	8.440
受入数	23.308	0.780	0.588	3.565	0.002	10.067
資料費	0.003	0.259	0.150	0.745	0.464	25.480
移動図書館の有無	-0.827	-0.086	-0.201	-1.006	0.324	1.534
移動図書館数	除去					
図書館面積	0.000	-0.171	-0.124	-0.613	0.545	16.404
図書館数	0.026	0.037	0.042	0.206	0.838	6.570
図書館密度	4.319	0.263	0.331	1.718	0.099	4.922
雑誌種数	449.090	0.312	0.419	2.261	0.033	3.991
開館時間	0.000	0.220	0.344	1.798	0.085	3.154
職員1000人あたり人口	0.000	0.313	0.295	1.513	0.143	8.989
重相関係数	0.941					
決定係数	0.886					
自由度調整済み決定係数	0.805					



の割合の影響力も高い点は共通であるが、専門職・管理職従事者の割合の影響力は限定的で、むしろ第1次産業就業者の割合の方が影響が大きかった。もっとも、個々の要因の影響力の高低は春日部（1990）でも年によって大きく変動することが示されており、また、多重共線性の目安となる VIF も、図書館内要因についていずれもかなり高い値を示しており（資料費で 25.480、図書館面積で 16.404、受入数で 10.067）、自治体に関する要因でも昼間人口（13.131）など高い値を示す変数がある。よって今回の説明変数の傾向の違いが、先行研究時点からの東京都の市区の変化であるとは位置づけられないであろう。岸田（1986）の検証同様、先行研究から長い期間を経た現在でも、回帰モデルによる貸出密度の説明は当時とほとんど変わらない精度で実現可能である、というのが本項で押さえるべき点と考えられる。

### 3. 岸田・佐藤（1991）：大阪府と富山県の市町

#### a. 大阪府

第4表に岸田・佐藤（1991）<sup>1)</sup> 同様、大阪府の市と町を対象に、変数減少法を用いて構築した、貸出密度を説明する重回帰モデルを示す。変数除去のための F 値確率は 0.10 とした。対象となった自治体数は 34 市町であった（岸田・佐藤（1991）では 1985 年度 33 市町、1980 年度 28 市町）。なお、春日部（1990）同様、岸田・佐藤（1991）についても専門職従事者と管理職従事者の割合の合算値を説明変数として採用している

ので、本項のモデルでも合算値を投入している。

最終的には 8 つの説明変数で、重相関係数 0.922、決定係数 0.851、自由度調整済み決定係数 0.803 のモデルが構築された。岸田・佐藤（1991）では、1980 年度データ使用の場合、10 の説明変数を用い、重相関係数 0.9326、決定係数 0.8697、自由度調整済み決定係数 0.7931、1985 年度データの場合は 5 つの説明変数を用い、重相関係数 0.8897、決定係数 0.7915、自由度調整済み決定係数 0.7529 のモデルが構築されていた。岸田・佐藤（1991）に比べても、ほとんど説明力の劣らないモデルが構築されたといえる。

採用された 8 つの説明変数のうち、特に影響力が強いのは職員 1000 人あたり人口、図書館面積、資料費と、いずれも図書館内要因である。自治体の状況に関する要因は女性の割合のみであった。また、岸田・佐藤（1991）のモデルとは採用される変数は大きく異なった。岸田・佐藤（1991）のモデルで 1980 年度データを用いた場合と、本研究の場合で重複しているのは資料費、図書館面積のみであり、1985 年度データ使用時との重複は資料費、受入数のみであった。それにも拘らず説明力に大きく差のないモデルが構築されているのは、変数間の多重共線性の問題が一因と考えられる。特に図書館内要因について、図書館面積は 9.180、職員 1000 人あたり人口は 10.403 と高い VIF を示しており、変数間に高い相関があることで採用される変数が安定しないのであろう。やはりここまでの結果同様、自治体側の要因と図

第4表 岸田・佐藤（1991）貸出密度予測再検証モデル・大阪府（大阪府の市町 34 自治体対象・変数減少法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
女性の割合	70.743	0.194	0.419	2.310	0.029	1.179
受入数	4.448	0.153	0.325	1.718	0.098	1.328
資料費	0.017	0.624	0.801	6.688	0.000	1.459
移動図書館数	1.203	0.305	0.540	3.211	0.004	1.513
図書館面積	0.000	1.111	0.688	4.743	0.000	9.180
図書館密度	-6.080	-0.169	-0.370	-1.994	0.057	1.198
開館時間	0.000	0.163	0.360	1.927	0.065	1.201
職員 1000 人あたり人口	0.000	-1.122	-0.669	-4.499	0.000	10.403
重相関係数	0.922					
決定係数	0.851					
自由度調整済み決定係数	0.803					

書館側の要因を用いて、高い説明力を持つモデルが構築できた、という点を押さえれば十分であろう。

#### b. 富山県

富山県について、当初は前項と同じく岸田・佐藤 (1991)<sup>1)</sup> 同様、富山県の市と町を対象に、変数減少法を用いて貸出密度を説明する重回帰モデルの構築を試みた。しかし変数減少法を用いた場合、重相関係数 1.000、決定係数 1.000 で、貸出密度と予測値がすべての図書館について完全に一致するモデルが構築されてしまった。これは明らかに過剰適応したモデルと考えられる。採用された各説明変数の VIF も大半が 10 以上、中には 50～60 を超える場合もあった。ここから富山県について、変数減少法によって貸出密度を説明する重回帰モデルを構築することは不適切と考えられる。

一方、岸田 (1986)<sup>3)</sup> と同様にステップワイズ法を用いた場合には一応のモデルを構築することは可能である。第 5 表に富山県の市と町を対象に、ステップワイズ法を用いて構築した、貸出密度を説明する重回帰モデルを示す。対象となった自治体数は 14 市町であった (岸田・佐藤 (1991) では 1985 年度 35 市町、1980 年度 26 市町)。先行研究に比べ自治体数が少ないが、これは『日本の図書館』の調査に回答している自治体が減ったり、開館時間や児童サービスに関する項目に未回答の自治体があるわけではない (富山県内の全自治体 10 市 4 町 1 村が『日本の図書館』2018 年版調査に回答しており、開館時間・児童サービスの項目にも回答している。なお 1 村は、村であるので先行研究に揃え、対象とはしていない)。そもそも市町村合併により、富山県内の自治体数が

減少していることが対象自治体数減少の要因である。なお、本項のモデルにおいても、専門職従事者・管理職従事者については合算値を説明変数として投入している。

最終的なモデルには 1 つの説明変数しか採用されず、重相関係数 0.559、決定係数 0.313、自由度調整済み決定係数 0.255 のモデルが構築された。採用された説明変数は蔵書数であり、実際にはこれは蔵書数で貸出密度を説明する単回帰モデルとなっている。

岸田・佐藤 (1991) では、1980 年度データ使用の場合、8 の説明変数を用い、重相関係数 0.8881、決定係数 0.7887、自由度調整済み決定係数 0.6892、1985 年度データの場合は 4 つの説明変数を用い、重相関係数 0.7060、決定係数 0.4984、自由度調整済み決定係数 0.3909 のモデルが構築されていた。岸田・佐藤 (1991) でも、1985 年度の富山県は数少ない、高い説明力を持つモデルが構築できなかった例として注目されていたが、本研究においては変数減少法を用いた場合にはそもそも有効なモデルの構築が困難であり、ステップワイズ法を用いた場合には顕著に説明力の劣るモデルしか構築できなかった。市町村合併による自治体数の減少が、自治体としての各指標の値と、利用実態の間に大きな乖離を引き起こした可能性がある (市町村合併の影響が少なかったであろう、首都圏や大阪府ではこのような変化は見られない)。

#### 4. 先行研究の再検証結果小括

本節では岸田 (1986)<sup>3)</sup>、春日部 (1990)<sup>11)</sup>、岸田・佐藤 (1991)<sup>1)</sup> の 3 つの先行研究が対象としていた自治体について、先行研究同様に説明力の高い貸出密度の予測モデルが現在でも構築できる

第 5 表 岸田・佐藤 (1991) 貸出密度予測再検証モデル・富山県 (富山県の市町 14 自治体対象・ステップワイズ法)

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
蔵書数	0.306	0.559	0.559	2.337	0.038	1.000
重相関係数	0.559					
決定係数	0.313					
自由度調整済み決定係数	0.255					

かを検証してきた。その結果、首都圏や大阪府といった大都市圏においては、先行研究同様に図書館に関する要因と自治体状況に関する要因を用いて、先行研究当時とほとんど変わらない説明力のモデルが、現在でも構築可能であることが確認された。具体的に影響力の高い説明変数は、先行研究当時とは異なる場合がほとんどであったが、これについては先行研究自体においても必ずしも一定しないことが指摘されており、その事実も含めて先行研究当時と同様であることが確認できた、と言えよう。

一方、富山県については先行研究時点でも、必ずしも説明力の高いモデルが構築できないことが報告されていたが、本研究ではその先行研究のモデルすら下回る説明力のモデルしか構築できないか（ステップワイズ法採用時）、あるいはそもそも有効なモデルが構築できなかった（変数減少法採用時）。市町村合併による自治体数の大幅な減少により、先行研究当時とは状況が一変したことも一因であろうが、「地方都市に対しては」説明力の高いモデルが構築できないのではないか、という先行研究時点での指摘の蓋然性は高まったと考えられる。次節以降ではこの指摘を念頭に置きつつ、全国の自治体を対象としたモデルの構築可能性と、地域を限定したモデルの構築可能性を検

証していく。

## B. 全国の自治体を対象としたモデルの構築

### 1. 全自治体対象モデル

第6表は日本のすべての市区町村（ただし『日本の図書館統計と名簿』2018年版に回答している、すなわち図書館が存在する、かつ前節同様、開館日数と児童への貸出冊数が0以外の自治体。以下同様）を対象に、ステップワイズ法を用いて構築した、貸出密度を説明する重回帰モデルである。対象となった自治体数は1,179市区町村であった。なお説明変数について、専門職従事者の割合と管理職従事者の割合は分けて投入している（以下同様）。

第6表のモデルにおいて重相関係数は0.705、自由度調整済み決定係数は0.491と、首都圏や東京都、大阪府、富山県に限定した先行研究のモデルや、前節で示したその再検証（ただし富山県を除く）に比べて低い説明力となっている。先行研究では「対象の質ならびに条件を揃える」ことを目的に地域や自治体種別を限定していたが、その方針は正しかったと言えよう。ただし、富山県を対象とする再検証結果に比べると第6表のモデルの方が説明力が高くなっている点には注意がいる。

第6表 日本の市区町村全体を対象とする貸出密度予測モデル（1,179自治体対象・ステップワイズ法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
雑誌種数	254.475	0.313	0.295	10.537	0.000	2.048
専門職従事者の割合	19.840	0.257	0.253	8.947	0.000	1.904
資料費	0.001	0.177	0.184	6.390	0.000	1.779
第1次産業就業者割合	-0.061	-0.212	-0.216	-7.566	0.000	1.822
非労働者の割合	-8.994	-0.155	-0.190	-6.604	0.000	1.275
図書館密度	4.402	0.145	0.163	5.638	0.000	1.522
蔵書数	0.115	0.172	0.183	6.364	0.000	1.692
開館時間	0.000	0.061	0.077	2.646	0.008	1.245
職員1000人あたり人口	0.000	-0.173	-0.169	-5.845	0.000	2.036
図書館面積	0.000	0.248	0.166	5.741	0.000	4.326
図書館数	-0.085	-0.099	-0.082	-2.805	0.005	2.870
管理職従事者の割合	-20.520	-0.050	-0.065	-2.212	0.027	1.177
重相関係数	0.705					
決定係数	0.496					
自由度調整済み決定係数	0.491					

説明変数の中でもっとも影響が大きいのは雑誌種数で、他には専門職従事者の割合、図書館面積、第1次産業就業者の割合（負の影響）も影響が大きい。説明力は必ずしも高くはないものの、説明変数として自治体側の要因、図書館側の要因の双方が影響力を持っている点は先行研究や前節までのモデルと同様と言える。

## 2. 市区対象モデル

第7表は日本の市区を対象に、ステップワイズ法を用いて構築した、貸出密度を説明する重回帰モデルである。対象となった自治体数は733市区であった。

重相関係数は0.765、自由度調整済み決定係数は0.579と、自治体種別を限定しないモデルに比べると説明力は高いものの、都市圏で地域と自治体種別を限定した場合に比べれば低くなっている。

説明変数の中でもっとも影響力が大きいのは第1次産業就業者の割合（負の影響）で、次いで専門職従事者の割合である。次が蔵書数であり、地域を限定した場合に比べ、自治体側の要因の影響が大きい傾向がうかがえる。

また、先行研究の中では地域を限定しないモデルを構築する場合、人口100万人以上の大都市を除くことが多い。本研究でもそれにならい、人口100万人以上の自治体を除いてモデルを構築し

た場合が第8表である。対象となった自治体数は723市区であった。

人口100万人以上の自治体を除いた場合であっても、重相関係数、自由度調整済み決定係数、採用される説明変数やその影響力の多寡など、いずれも大きな変化はなかった。超大規模自治体であっても特に例外的なふるまいを示しているわけではなく、他の市区同様に分析に組み込んでも問題はないと考えられる。

## 3. 町対象モデル

第9表は日本の町を対象に、ステップワイズ法を用いて構築した、貸出密度を説明する重回帰モデルである。対象となった自治体数は410町であった。

重相関係数は0.703、自由度調整済み決定係数は0.486で、第6表の自治体種別を限定しないモデルよりも説明力が低いモデルしか構築できなかった。説明変数の中では資料費と図書館面積の影響がほかに差をつけて大きく、その後、専門職従事者の割合や第1次産業就業者の割合等の自治体側の要因が出てくる。町に限定すると、自治体側の要因よりも図書館側の要因の影響が貸出密度に対しては大きいと言える。

## 4. 村対象モデル

第10表は日本の村を対象に、ステップワイズ

第7表 日本の市区を対象とする貸出密度予測モデル（733自治体対象・ステップワイズ法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
雑誌種数	303.148	0.193	0.204	5.611	0.000	2.048
第1次産業就業者割合	-0.101	-0.271	-0.319	-9.029	0.000	1.563
職員数	2.546	0.199	0.210	5.761	0.000	2.072
専門職従事者の割合	17.475	0.243	0.276	7.728	0.000	1.725
非労働者の割合	-7.464	-0.131	-0.172	-4.703	0.000	1.345
蔵書数	0.379	0.242	0.241	6.675	0.000	2.295
人口密度	0.000	0.136	0.156	4.257	0.000	1.784
資料費	0.001	0.107	0.150	4.087	0.000	1.197
開館時間	0.000	0.095	0.131	3.548	0.000	1.258
管理職従事者の割合	-29.678	-0.079	-0.105	-2.838	0.005	1.342
重相関係数	0.765					
決定係数	0.585					
自由度調整済み決定係数	0.579					

公共図書館における貸出関数の可用性の再検証

第8表 日本の市区（人口100万人以上を除く）を対象とする貸出密度予測モデル（723自治体対象・ステップワイズ法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
雑誌種数	297.812	0.189	0.202	5.509	0.000	2.032
第1次産業就業者割合	-0.096	-0.256	-0.304	-8.494	0.000	1.579
職員数	2.460	0.191	0.203	5.521	0.000	2.071
専門職従事者の割合	17.168	0.237	0.271	7.493	0.000	1.733
非労働者の割合	-7.224	-0.126	-0.167	-4.527	0.000	1.341
蔵書数	0.380	0.240	0.243	6.672	0.000	2.241
人口密度	0.000	0.125	0.142	3.834	0.000	1.837
資料費	0.001	0.104	0.147	3.965	0.000	1.194
開館時間	0.000	0.086	0.117	3.139	0.002	1.290
管理職従事者の割合	-30.473	-0.081	-0.108	-2.909	0.004	1.343
図書館面積	0.000	0.067	0.091	2.426	0.016	1.338
重相関係数	0.768					
決定係数	0.590					
自由度調整済み決定係数	0.584					

第9表 日本の町を対象とする貸出密度予測モデル（410自治体対象・ステップワイズ法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
資料費	0.004	0.443	0.464	10.516	0.000	1.414
図書館面積	0.001	0.320	0.388	8.451	0.000	1.143
第1次産業就業者割合	-0.051	-0.204	-0.212	-4.341	0.000	1.765
専門職従事者の割合	17.576	0.205	0.217	4.453	0.000	1.686
昼間人口	2.061	0.091	0.122	2.466	0.014	1.089
蔵書数	0.054	0.107	0.127	2.564	0.011	1.379
非労働者の割合	-4.649	-0.079	-0.104	-2.092	0.037	1.139
重相関係数	0.703					
決定係数	0.495					
自由度調整済み決定係数	0.486					

第10表 日本の村を対象とする貸出密度予測モデル（36自治体対象・ステップワイズ法）

変数	回帰係数	標準回帰係数	偏相関係数	t	p	VIF
図書館密度	54.169	0.591	0.693	5.516	0.000	1.231
雑誌種数	226.460	0.384	0.529	3.584	0.001	1.231
重相関係数	0.832					
決定係数	0.693					
自由度調整済み決定係数	0.674					

法を用いて構築した、貸出密度を説明する重回帰モデルである。対象となった自治体数は36村であった。

重相関係数0.832、自由度調整済み決定係数0.674と、市区や町を対象とするモデルに比べて高い説明力を持つモデルが構築できた。図書館が

ある村自体、数が少ないことも一因と考えられるが、それらの図書館を持つ村は、市区や町に比べ、貸出密度に影響する何らかの要因の均質性が高いと考えられる。

説明変数の中でもっとも影響が大きいのは図書館密度であり、それと雑誌種数の2つの図書館側



の要因のみが採用された。ただし、実際には2つ以上の図書館を持つ村はないため、図書館密度は実質的に、村の面積の逆数であり、自治体側の要因ととらえることができる。

### C. 地域を限定したモデルの構築

8つの地域別に、貸出密度を説明する重回帰モデルを構築した結果のまとめを第11表、第12表に示す。第11表は地域ごとのモデルに含まれる自治体数、説明変数の数、重相関係数、決定係数、自由度調整済み決定係数のまとめ、第12表はモデルに採用された説明変数と回帰係数、標準回帰係数、偏相関係数のまとめである。

表のとおり、説明力の高いモデルが構築できるかどうかは地域によって明確に異なる。ただし、具体的にどのような傾向があるかは、先行研究とその再検証から得られていた推測（「地方都市に対しては説明力の高いモデルが構築できない」）とは大きく異なる。第11表のとおり、関東においては、先行研究のように首都圏（埼玉県、千葉県、神奈川県、東京都）に絞った場合に比べれば落ちるものの、十分に説明力の高いモデルが構築できている。これを大都市圏においては説明力の高いモデルが構築できる、という推測の裏付けとしたくなるが、実際には関東よりも説明力の高いモデルが構築ができている地域として、四国がある。重相関係数0.920、自由度調整済み決定係数0.812というのは先行研究と比べても最も説明力の高いモデルである。また、関東に次いで3番目に説明力の高いモデルが構築できたのも東北地方である。東北地方には仙台市が存在するといっても、大阪市や京都市、神戸市を擁する近畿と比べて大都市圏であると言い難いであろう。

では大都市圏より地方の方が説明力が高いモデ

ルが構築できるのかといえ、そのようなわけでもない。最も説明力の低いモデルしか構築できなかった北海道（自由度調整済み決定係数は0.186）は、自治体数の少なさも一因と言えるかもしれないが、広島市が存在する以外は大都市圏とはいえない中国、福岡市以外は同じく大都市圏と言い難い九州なども、説明力の高いモデルは構築できていない。このように、地方都市であるか、大都市圏であるかという観点では、地域ごとの差を捉えることはできないといえよう。

モデルに採用される説明変数についても地域によって異なるが、図書館側の要因としては雑誌種数（5地域で採用）、蔵書数（4地域で採用）、職員数（4地域で採用）、開館時間（4地域で採用）、自治体側の要因としては女性の割合（4地域で採用）が採用されることが多かった。図書館側の要因としては移動図書館の有無・台数、自治体側の要因としては管理職従事者の割合と第3次産業就業者割合は全く採用されなかった（ただし第3次産業就業者割合については、第1・2次産業就業者割合と完全に連動しているためと考えられる）。

## V. 考察

本稿では自治体を単位とする線形回帰モデルによる貸出関数に着目し、その可用性の検証を目的とし、以下のリサーチ・クエスチョン（RQ）の下でデータの分析を行った結果を報告してきた。

RQ1: 先行研究で対象とされていた地域において、現在でも線形回帰モデルにより、高い決定係数を持つ貸出密度の予測モデルが構築できるか

RQ2: 先行研究で対象とされていない地域において、線形回帰モデルにより、高い決定係数を持つ

第11表 地域ごとの貸出密度予測モデルに含まれる自治体数、変数、決定係数等のまとめ

	北海道	東北	関東	中部	近畿	中国	四国	九州
自治体数	28	67	179	151	119	49	34	106
説明変数の数	1	5	8	5	7	3	6	5
重相関係数	0.465	0.798	0.867	0.673	0.782	0.658	0.920	0.707
決定係数	0.217	0.637	0.753	0.453	0.611	0.432	0.847	0.500
自由度調整済み決定係数	0.186	0.608	0.741	0.435	0.586	0.395	0.812	0.474

第 12 表 地域ごとの貸出密度予測モデルに含まれる説明変数と回帰係数等のまとめ<sup>1)</sup>

変数	北海道				東北				関東				中部			
	回帰 係数	標準回帰 係数	偏相関 係数	VIF	回帰 係数	標準回帰 係数	偏相関 係数	VIF	回帰 係数	標準回帰 係数	偏相関 係数	VIF	回帰 係数	標準回帰 係数	偏相関 係数	VIF
女性の割合	-70.655	-0.465	-0.465	1.000	-35.405	-0.242	-0.355	1.122	12.538	0.195	0.300	1.546	-26.015	-0.167	-0.192	1.335
非労働者の割合																
専門職従事者									-0.442	-0.181	-0.263	1.788				
管理職従事者					0.001	0.315	0.443	1.124	0.263	0.149	0.186	2.523	0.000	0.185	0.208	1.383
第 1 次産業割合					6.155	0.325	0.427	1.312	9.865	0.280	0.320	2.772				
第 2 次産業割合													0.002	0.227	0.243	1.495
第 3 次産業割合					2.409	0.211	0.276	1.480	4.683	0.370	0.358	3.782				
昼間人口																
人口密度																
蔵書数																
受入数																
資料費																
職員数																
移動図書館の有無																
移動図書館数																
図書館面積									0.056	0.092	0.155	1.379				
図書館数																
図書館密度									200.748	0.119	0.162	2.112	704.654	0.424	0.417	1.557
雑誌種数					0.000	0.313	0.444	1.103	0.000	0.107	0.178	1.407	0.000	0.218	0.276	1.053
開館時間																
職員 1000 人あたり人口																

第12表 続き

変数	近畿				中国				四国				九州			
	回帰 係数	標準回帰 係数	偏相関 係数	VIF	回帰 係数	標準回帰 係数	偏相関 係数	VIF	回帰 係数	標準回帰 係数	偏相関 係数	VIF	回帰 係数	標準回帰 係数	偏相関 係数	VIF
女性の割合																
非労働者の割合																
専門職従事者	26.670	0.360	0.435	1.431												
管理職従事者																
第1次産業割合	-0.092	-0.184	-0.236	1.480												
第2次産業割合																
第3次産業割合																
昼間人口																
人口密度																
蔵書数	0.475	0.354	0.332	2.605												
受入数	-1.714	-0.183	-0.257	1.228												
資料費																
職員数	6.048	0.324	0.336	2.117	2.047	0.261	0.317	1.080								
移動図書館の有無																
移動図書館数																
図書館面積					0.000	0.474	0.529	1.021								
図書館数																
図書館密度																
雑誌種数	297.363	0.185	0.185	2.482	479.617	0.392	0.449	1.069								
開館時間	0.000	0.164	0.228	1.263												
職員1000人あたり人口																
					0.000	0.379	0.654	1.254								

1) t 値, p 値は省略している。また,  $p < 0.01$  の変数には二重下線,  $p < 0.05$  の変数には一重下線を各値に付与している。

貸出密度の予測モデルは構築できるか

RQ2-1: 地域を限定しないモデルの構築は可能か

RQ2-2: 地域を限定したモデルの場合、地域ごとに妥当なモデルは構築できるか

以下ではまずこれらの RQ について分析結果に基づき論じていった後、全体として線形回帰モデルによる貸出関数の可用性をどう捉えるべきか、今後どういった点の検証が必要かを述べる。

#### A. RQ1: 先行研究対象地域における線形回帰モデルによる貸出関数の可用性

前章の小括でも示したとおり、先行研究で線形回帰モデルを用いた貸出密度の予測の対象としていた地域のうち、首都圏、東京都、大阪府といった大都市圏においては現在においても、先行研究が対象としていた約 30 年前とほとんど変わらない説明力のモデルが構築可能である。こういった説明変数（図書館・自治体の要因）の影響力が大きいかは対象とする地域によって異なったが、これは先行研究当時から指摘されていたことである（ステップワイズ法や変数減少法においてこういった説明変数が採用されるか、強制投入法においてどの説明変数の影響が大きいかは地域や時期により一定しない。これは多重共線性の影響によると考えられる）。そのため具体的に図書館や自治体のどの要因が貸出密度に影響するかをはっきりと特定することは困難であるものの、それぞれの要因が貸出密度の予測に役立つことは確かであり、極論すれば強制投入法ですべて投入してしまえば、ある程度の予測は実現可能といえる。

ただし、先行研究で対象とした地域の中でも富山県については、先行研究当時でも必ずしも説明力の高いモデルが構築できていなかったとはいえ、本研究においてはそれよりさらに説明力の低いモデルしか構築できず、実用可能なレベルでの貸出密度の予測は実現できなかった。市町村合併による自治体数自体の減少も大きな要因として考えられるが、先行研究において提示されていた、大都市圏においては説明力の高いモデルが構築できるが、地方都市においては構築できないのでは

ないか、という仮説の蓋然性が高まった、とこの段階では考えられた。しかしこの仮説は後述のとおり、先行研究対象地域以外におけるモデル構築の試みにおいて、棄却されることとなる。

#### B. RQ2: 先行研究で対象とされていない地域、あるいは地域を限定しない線形回帰モデルによる貸出関数の可用性

まず日本の全自治体（市区町村）を対象とした場合（RQ2-1）には、地域を絞った先行研究やその再検証の結果（富山県以外）に比べて説明力の低いモデルしか構築できなかった。この時点では「対象の質ならびに条件を揃える」ために地域や自治体種別を限定する先行研究の方針は正しかったのではないかと考えられた。

しかし自治体種別を絞って分析してみると、確かに市区に限定すれば限定しないよりも説明力の高いモデルが構築できるものの、町に限定した場合には限定しないよりも説明力は低くなり、一方で村に限定すると市区や、先行研究等における一部の大都市圏（大阪府）よりも説明力の高いモデルが構築できるなど、一貫しない結果となった。

さらに 8 つの地域別に自治体種別を絞って分析してみると（RQ2-2）、この一貫性のなさはより明白となった。最も説明力の高いモデルが構築できたのは政令指定都市が一つもない四国、二番目に説明力の高いモデルが構築されたのは大都市圏を含む関東、しかし三番目は仙台市があるとはいえ都市圏とはいえない東北、となった。この結果を見る限り、説明力が高いモデルが「地方都市においては構築できないのでは」とする先行研究の指摘はあてはまらないと言えよう。

そもそも地域を限定した場合に、地域を限定しないモデルよりも高い説明力を示した地域自体、決して多くはない。地域を限定せず、自治体種別を市区に限定した全国モデルの自由度調整済み決定係数は 0.579 であるが、これを超える説明力を示した地域限定モデルは東北、関東、近畿、四国でしか構築されていない。さらに言えば、地域を限定したモデルのうち、先行研究やその再検証で高い説明力を示した「首都圏の市区」や「東京都

の市区」を超えるか匹敵する説明力のモデルが構築されたのは、四国のみであった。その点では、そもそも貸出密度の予測モデル、ひいては貸出関数の予測において、「地域を限定する」という手続き自体の妥当性が疑問視される。これについては次節で詳しく論じていきたい。

### C. 図書館・自治体の要因を用いた貸出関数の可能性と、それを高めるために今後必要な検証：「要因 X」の特定

RQ1 から、先行研究で対象としていた地域においては現在でも線形回帰モデルによる貸出密度の予測が可能であることが示された。ここから、図書館や自治体の要因を用い、貸出関数を構築する、というアプローチ自体は有望であると考えられる。しかし RQ2-1 から、単純に同じモデルを全国規模に広げることができないことが示された。そこまでは予測のとおりであったが、RQ2-2 の地域別の分析からはさらに、先行研究対象地域以外では、地域を限定したからといって線形回帰モデルによる貸出密度の予測が高精度で実現できるわけではない、ということも示された。さらに、どういった地域で貸出密度の予測が高精度で実現できるかは、必ずしも大都市圏か否か等だけでは説明できないこともわかった。

最初期の先行研究では首都圏、あるいは東京都のみを対象に貸出密度の予測モデルを構築し、モデルに含んだ説明変数である程度、貸出密度を予測できることを示したが、本研究の結果から、これはある意味、「あたり」を最初に引き当てたものであったと言える。他の地域では必ずしも同程度の説明力を持つモデルが構築できるとは限らない。しかも、それは必ずしもその地域の都市化の程度等とは関連していない（四国が最も説明力の高いモデルが構築できる）。

なぜこのような結果になるのかは、ある意味ではすでに先行研究で言及されている。そもそも先行研究で首都圏等に地域を限定したのは“対象の質ならびに条件を揃える”<sup>3)</sup> ためである（ここでいう対象は自治体）。これは言い換えれば、モデル内に含まれうる説明変数の候補以外に、なんら

かの貸出密度に予測する要因が存在すると考え、その要因を地域を限定することで統制できると考えた、ということである。そして実際に首都圏等の内部においてはそのなんらかの要因が自治体ごとに似通っていたために、地域を限定することで説明力の高いモデルが構築できたものと考えられる。以下、このなんらかの要因のことを「要因 X」と呼称する（実際には2つ以上の要因があることも考えられるが、暫定的に一つの要因として扱う）。最初期に幸運にも、要因 X が統制されている地域が対象に選ばれたため、その後、初期のモデル内に含まれる変数で（多重共線性に目をつぶれば）貸出の予測はある程度可能、とみなされることになってしまったが、実際には本研究で見たとおり、要因 X が統制されていない状況では十分な説明力を持つモデルが構築できない場合も多く、かつ地域を限定したからといって必ずしも要因 X も統制できるわけではない。よって要因 X が特定できなければ、一部地域以外では実用可能な貸出関数のモデルは構築できないと言える。逆に要因 X を特定することができ、かつそれがなんらかの形で計測・入手可能なデータであったとすれば、それを組み込むことで貸出関数の可用性はさらに高まり、一部地域以外でも、さらには地域を特定せずとも実用可能なものになることが考えられる。

では要因 X とはなんなのか。単純な地理的近縁性（例えば近隣地域の図書館の利用状況等は、自治体や図書館側の要因に依らず似通る、など）であったとすれば、首都圏以外でも地域を限定すれば予測モデルの説明力は向上するはずである。しかし本研究の結果から、実際には地域を限定しても説明力が向上しないケースも多いことが示されており、要因 X が地理的近縁性である可能性はすでに否定された。

要因 X のその他の候補を考える上でも、本研究の結果が参考となるだろう。すなわち、「四国」という区切りで自治体間の均質性が最も高く、「東京都」、「首都圏」、「関東」といった区切りでも高い、また「東北」や「近畿」という区切りでもある程度、均質になるが、「中国」、「北海



道」といった区切りでは均質性が低くなる要因とは何か、ということを考えていくことが、要因X特定の鍵になると考えられる。その要因が図書館側の要因か、自治体側の要因かは現時点では定かではないものの、図書館側の要因として考えられるものは既にある程度、モデル内に含まれていることを踏まえると、自治体側の要因である蓋然性が高い。とはいえ、この要因Xが実は「図書館の努力」であり、要因Xを導入しないモデルにおける予測値と実測値の差こそが、個々の図書館の、蔵書数や職員数等の数にあらわれないサービスの「質」なのだ、という可能性も皆無ではない。その可能性は、他にどれだけ考えられる要因を検証しても要因Xが特定できなかった場合に、より現実味を帯びてくることとなる。

本稿では紙幅の都合もあることから、具体的に要因Xを特定する作業（先行研究には含まれていなかった説明変数の候補を投入し、より説明力の高いモデルが構築できるかを検証する作業）については今後の課題としたい。しかし「はじめに」で述べた公共図書館評価の新たな在り方（図書館利用の標準値を算出し、実測値との乖離状況を測り、その理由を検討する）を実現していく上では、まずはこの要因Xの特定が急務であり、迅速に取り組んでいきたい。

最後に要因Xが実際に特定され、実用に耐える貸出関数が利用可能になった場合にどのような図書館評価が可能になるか、第一著者自身の経験に基づく一例を述べたい。第一著者自身がある公共図書館の事業計画策定に協力する中で、当該図書館の従来の活動に関する統計データ等は当然、提示されたものの、それがあつ程度、妥当な数値なのか、一層の努力が必要な値なのか簡単には判断できない、という状況に見舞われた。当該自治体は近隣では第二位の人口規模を有し、専任職員数も多いが、図書館の規模は自治体規模に比して小さく、また当面、追加で予算が付くことも見込めない。その中で例えば貸出密度の値は周辺自治体に比べて必ずしも高くはないが、ではこの値は現在の自治体・図書館の状況に照らして妥当な（標準的な）値なのか、それとも「努力不足」と

受け止めるべき値なのか。判断の参考になる指標がないかと考えた際に、本研究で構築した貸出関数の予測値と比較することを思いついた。実際に見比べたところ、当該自治体の貸出密度はほぼ予測値と一致していた。貸出のみについてはあるが、当該図書館は標準的な状況にあると判断し、それに基づいて事業計画への助言を行うことができた。現状の貸出関数はまだ十分な予測精度を持つものではなく、各図書館の標準値・実測値を一般に公開すべき段階とは考えていないが、今後、要因Xの検証作業が十分に進んだ際には、各図書館の標準値・実測値を公開し、多くの人々が上述のような際に参考にできるようにすることを目指す。

## 謝 辞

本研究はJSPS科学研究費補助金19H04428の支援を受け行いました。

## 注・引用文献

- 1) 岸田和明、佐藤佳子. 公共図書館の貸出を説明する関数の重回帰分析による検証：大阪府および富山県を対象として. *Library and Information Science*. 1991, no. 29, p. 161-168.
- 2) 田辺智子. 図書館評価と行政評価：評価方法の共通点と相違点の分析. *図書館情報メディア研究*. 2016, vol. 14, no. 1, p. 21-39.
- 3) 岸田和明. 公共図書館の利用に影響を与える要因. *Library and Information Science*. 1986, no. 24, p. 45-55.
- 4) Kim, C.; Shin, E. H. Sociodemographic correlates of intercounty variations in the public library output. *Journal of the American Society for Information Science*. 1977, vol. 28, no. 6, p. 359-365.
- 5) 糸賀雅児. 公共図書館の活動指標と図書館内的要因の分析. *図書館学会年報*. 1982, vol. 28, no. 1, p. 13-28.
- 6) 常盤繁. 重回帰分析による公共図書館利用量（個人貸出冊数）の分析. *情報科学研究*. 1985, no. 3, p. 15-30.
- 7) Parker, E. B.; Paisley, W. J. Predicting library circulation from community characteristics. *Public Opinion Quarterly*. 1965, vol. 29, no. 1, p. 39-53.
- 8) Carlozzi, M. The socioeconomic profile of well-funded public libraries: A regression analysis. *Evidence Based Library and Information Prac-*

- tice. 2018, vol. 13, no. 2, p. 13-26.
- 9) Seavey, C. A. Public library systems in Wisconsin, 1970-1980; An evaluation. The University of Wisconsin, 1987, Ph.D. thesis.
  - 10) Koontz, C. M. Public library site evaluation and location: Past and present market-based modeling tools for the future. Library & Information Science Research. 1992, vol. 14, no. 4, p. 379-409.
  - 11) 春日部千絵. 公共図書館の利用要因の時系列的分析. Library and Information Science. 1990, no. 28, p. 121-143.
  - 12) 岸田和明. 規模要因を考慮した公共図書館の貸出に関する数量的モデル. Library and Information Science. 1998, no. 40, p. 29-35.
  - 13) 岸田和明. “計量書誌学および地理的要因を考慮した公共図書館の活動に対する評価指標”. 三田図書館・情報学会研究大会発表論文集. 東京, 2013-11-09, 三田図書館・情報学会. 2013, p. 29-32.
  - 14) 田村肇. 公共図書館の貸出関数の実証的分析. 図書館情報大学研究報告. 1994, vol. 13, no. 1, p. 73-80.
  - 15) 田村肇. 経済状況と公共図書館の利用: 都道府県パネルデータを用いた分析. 日本図書館情報学会誌. 2004, vol. 50, no. 2, p. 58-78.
  - 16) 金井喜一郎. 公共図書館の貸出関数の検証. 大妻女子大学紀要, 社会情報系社会情報学研究. 2017, vol. 26, p. 107-114.
  - 17) 日本図書館協会図書館調査事業委員会. 日本の図書館調査委員会編. 日本の図書館: 統計と名簿 2019 [CD-OM 版]. 東京, 日本図書館協会, 2019.
  - 18) 総務省統計局調査部国勢統計課. “平成 27 年都道府県・市区町村別統計表 (国勢調査)”. e-Stat. 2017-06-28. [https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001049104&cycle=0&tclass1=000001049105&stat\\_infid=000031594311&tclass2val=0](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001049104&cycle=0&tclass1=000001049105&stat_infid=000031594311&tclass2val=0), (参照 2021-03-10).

## 要 旨

**【目的】** 貸出回数を被説明変数, それに影響を与えうる要因を説明変数とする「貸出関数」は, かつて日本の公共図書館を対象とする複数の研究が行われたものの, 近年では必ずしも盛んには研究されていない。本研究の目的は公共図書館の貸出関数の中でも, 自治体を単位とする線形回帰モデルによる貸出関数に着目し, その可用性を再検証することである。

**【方法】** 本研究では主な先行研究同様, 被説明変数として貸出密度 (人口あたりの年間貸出冊数) を採用し, 説明変数の候補としては蔵書数等の図書館内の要因と, 人口密度等の自治体自体の状況を投入した, 重回帰モデルを構築する。対象とする自治体は(1)先行研究で対象とされた地域 (首都圏の市等), (2)日本全国の市区町村, (3)日本の市区を 8つの地域に区分した場合, とする。(1)は先行研究が対象とした地域で現在でも高い説明力のモデルが構築できるかを検証すること, (2)・(3)は先行研究で対象としていない地域でも高い説明力のモデルが構築できるかを検証することを目的とする。

**【結果】** 先行研究で対象とされていた地域の多くにおいては, 自由度調整済み決定係数 0.7~0.8 と, 先行研究に匹敵する説明力のモデルが構築できた。一方, 日本全国を対象とした場合には説明力は大きく下がり, さらに先行研究対象地域以外では, 地域を分けても必ずしも説明力の高いモデルが構築できるわけではなかった。