

短 報

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

A Mathematical Model on Public Library Loan
Incorporating an Inverse Size Effect岸 田 和 明
Kazuaki Kishida

Résumé

The purpose of this paper is to propose a new mathematical model for predicting the amount of loan at public libraries, and to investigate the validity of the model by using statistical data on public libraries at Saitama Prefecture in Japan. The model is $L = aCPexp(-bC)$, where L is the number of times books are circulated in a library, C is the number of books held at the library, P is the population within the service area, and a and b are parameters. The model explains well statistical variation of the amount of loan at Saitama Prefecture. It can be said that the result shows a sort of "inverse effect" of size of collection or population, which is represented by $exp(-bC)$ in the model, on the increase in loan. Also, a simple simulation based on the model is tried to assess the effectiveness of new sites for access.

- I. はじめに
- II. 貸出モデルの演繹的導出
- III. データによる実証
- IV. モデルの解釈とシミュレーション
- V. おわりに

I. はじめに

公共図書館の貸出延べ冊数（貸出回数）に関するマクロ的な計量モデルは、一般に、

$$L = f(x_1, x_2, \dots, x_p) \quad (1)$$

と表わすことができる。ここで、 L はある図書館における一定期間の貸出延べ冊数、 x_1, x_2, \dots, x_p

はその貸出に影響を与える要因である。具体的に f として何らかの関数を設定してそのパラメータをデータから推定することにより、パラメータの大きさ等を通じて、図書館の貸出に関する特性や要因を分析することができる。また、このモデルによる貸出予測を公共図書館の建設計画や地域計画に役立てることも可能である。

岸田和明：駿河台大学文化情報学部，埼玉県飯能市阿須 698
Kazuaki Kishida: Surugadai University, Azu 698, Hanno-shi, Saitama-ken
受付日：2000年1月18日 受理日：2000年2月19日

このため(1)式のモデルに対して数多くの研究が積み重ねられてきた。そこでは、多くの場合、線型回帰モデルあるいは対数線型回帰モデルが利用されている。すなわち、

$$L_i = a_0 + a_1x_{i1} + a_2x_{i2} + \dots + a_px_{ip} + \varepsilon_i \quad (2)$$

あるいは、

$$\log L_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + \bar{\varepsilon}_i \quad (3)$$

であり、ここで L_i は図書館*i*における貸出延べ冊数、 x_{ij} ($j=1, \dots, p$)は、図書館*i*における要因*j*についての説明変数、 ε_i と $\bar{\varepsilon}_i$ は図書館*i*における誤差項(独自要因)、 $a_0, \dots, a_p, b_0, \dots, b_p$ はパラメータである。

実際にモデルには多種多様な説明変数が投入されているが、大きく分ければ、それらは、蔵書冊数や受入冊数などの図書館の内的な要因と、地域の人口や教育水準などの外的な要因とに区分できる。後者の外的な要因を組み込む場合には、通常その統計が地方自治体ごとに算出されるため、図書館ではなく、都道府県あるいは市区町村を単位とした分析がおこなわれる。

このような研究は欧米においていくつか試みられているが¹⁾、我が国においても、類似の分析例は数多い(例えば、糸賀²⁾や岸田³⁾、春日部⁴⁾、岸田・佐藤⁵⁾など)。また、図書館計画あるいは経営分析のために、回帰分析が試みられる例もある⁶⁾。さらに時系列的な要素を明示的に加えた研究例として田村⁷⁾がある。そこでの主な結果は、蔵書冊数や受入冊数などの説明変数間には高い内部相関があり、それらはまた、被説明変数である貸出延べ冊数とも高い相関があって、その変動をかなりの程度説明することが可能というものである。

以上のような(2)式あるいは(3)式を用いた研究の不十分な点の1つは、これらが理論的に導出されたモデルというよりも、データを従来の回帰分析などの統計的手法(あるいは統計パッケージ)に機械的に「ほうり込んだ」ものという印象を拭いきれない点にある。もちろん、上記の諸研究の中にはモデル構築に主たる関心があるわけではないものも含まれているし、また、回帰モデルを利用する場合にも、どのような説明変数を投入するかについては、理論的な根拠づけや検討がな

されるのが普通である。しかし、その理論的考察は、モデルの構造自体までには十分に反映されていない。実証的な統計データを駆使してさまざまな事実を明らかにし、公共図書館の貸出に関する知見を増やしたという点で、上に掲げた諸研究が評価されるのは妥当ではあるが、さらなる次のステップとして、(1)式についての探究が必要であると考えられる。

本稿では、(1)式に示されている f として、(2)式あるいは(3)式とは異なるモデルを演繹的に導出し、その妥当性を簡単な実証的データによって検証することを試みる。具体的には、まず第二章において演繹的モデルを導出し、第三章で、埼玉県市部の図書館データを使ったモデルの検証をおこなう。そして、第四章では、このモデルを利用した簡単なシミュレーションを試み、モデルの有用性を例示する。

II. 貸出モデルの演繹的導出

最初の基本的仮定として、筆者の以前の研究³⁾と同様に、ある公共図書館の貸出延べ冊数の多寡には、図書館の内的な要因だけではなく、図書館の外的な要因も影響すると考える。もし仮に同一の蔵書規模あるいはサービスの質を持つ図書館があった場合、そのサービス対象である人口規模が異なれば、その大きな方の図書館における貸出延べ冊数が他方のそれよりも多くなると考えるのは妥当であろう。内部的要因を I と表記し、外部的要因を E とかく。すなわち $L=f(I, E)$ である(なお、煩雑さを省くため、これ以降添え字*i*を省略する)。

さらに、 I, E はそれぞれ $I=g(y_1, \dots, y_m), E=h(z_1, \dots, z_n)$ のように数多くの他の要因の関数と捉えることができる。例えば、前者は蔵書冊数や受入冊数、開館時間、後者は定住人口や教育水準、立地条件などの関数として考えることはこれまでの研究成果^{1)~5)}からすれば妥当であろう。しかし、本研究では、これらの要因を詳細に取り上げるのではなく、前者を蔵書冊数 C 、後者を定住人口 P で代表させることにしたい。

これは、これまでの研究成果^{1)~5)}に照らせば、

y_1, \dots, y_m の間,あるいは z_1, \dots, z_n の間には,かなりの程度の内部相関が存在し,それぞれそのうちの1つを取り上げれば貸出延べ冊数の変動の「近似的な」説明には十分と考えられるためである。もちろん,開館時間や教育水準のように,直感的には,それぞれ蔵書冊数や定住人口との相関がそれほど高くないと予想される変数もある。しかし,そのような変数を投入したとしても,マクロなレベルで見れば,それぞれの独自の寄与は比較的小さく,蔵書冊数と定住人口のみを考慮するモデルが1つの近似として十分である可能性が高い。また,やや消極的な理由として,考慮する変数が多くなると,演繹的な数学モデルの構築が難しくなるということもある。

以上のような理由で,蔵書冊数と定住人口のみを説明変数とするモデルを1つの近似として,これらの変数のみを取り上げることの不十分さを認めた上で,そのモデルの構築を進めることとした。当然,本研究は,別の説明変数をも組み込んだ,より複雑なモデルの将来的な構築の可能性を否定するものではない。

さて,蔵書冊数と定住人口のみを取り上げた上で,次にこれらの間の関係を考察してみよう。仮に,蔵書が1冊で利用者が1人のとき,その人がそれを借りるとすると,貸出延べ冊数は1冊である。次に,蔵書2冊,利用者2人で,その2人とも両方借りるとすれば貸出延べ冊数は4冊になる。したがって, C と P は相乗の効果を持つと考えるのが妥当であろう。少なくともこの場合相加的な効果($C+P$)を想定するのは意味がない。例えば,蔵書冊数1冊で利用者1人の場合に,貸出延べ冊数を $1+1=2$ と計算する合理的な理由を見出すことは難しい。もちろん,すべての利用者がすべての蔵書を借り出すわけではないから,定数 a を掛けて aCP とすべきである($0 < a < 1$)。

しかし, C と P が大きくなれば, L も無制限に大きくなると考えるのは非現実的である。まず,地理的な要因がある⁸⁾。つまり,人口規模 P が大きくなるということは,人口密度には限界があるから,人々の居住範囲がある程度広がることを意味するが,アクセスポイントとしての図書館の数

をそれに応じて増やさない限り,居住場所から図書館までの距離が遠くなるので,貸出回数が単純に P に比例して大きくなるとは考えにくい。また,いたずらに C が大きくなっても,我が国の出版点数には限界があるので,その蔵書規模に応じた本(利用者の関心を惹く本)を量的に揃えることができることも限らない。

そこで,規模要因 S を導入して,それが大きさと逆の方向に作用するような「抑制」をモデルに加える必要がある。このためには,減衰の効果を表す非常に一般的な関数 e^{-x} を用いて, $aCPe^{-S}$ とすることがまず考えられる。実際に S をどのように実測するかは大きな問題であるが,ここでは,蔵書規模 C で代替させる(パラメータを付加して bC とする)。もちろん, P や他の変数,あるいはそれらの関数でも可能かもしれない。

以上の考察から,モデルは,結局,

$$L = aCPe^{-bC} \quad (4)$$

となる(a と b はパラメータ)。これが本研究で提案する最終的な貸出モデルである。

III. データによる実証

ここではモデル(4)式の妥当性を簡単なデータを用いて実証する。データとしては(社)日本図書館協会による『日本の図書館1994』⁹⁾を用いる。入力の手間・誤りを省くため,磁気テープ版を購入し,駿河台大学の大型計算機上でプログラムを作成して必要なデータを抽出した。その後,それをMS-DOSファイルに変換し,以下の処理・分析はすべてMicrosoft社のExcel for Windows 95,およびそれに付属するVBA(Visual Basic for Applications)でおこなった。

今回は,埼玉県の一部のみを対象とすることとした(全部で42市)。埼玉県というわずか1つの県に限定したのは,モデルの検証の第1段階としてまず対象地域を狭くしておき,万一モデルが不十分であった場合にその原因を細かく探究できるように配慮したためである。そしてさらに,市と町村ではその図書館や地域の状況・特性が多少異なっていると考え,性質の異なる2つの集団を混合して分析する危険を避けるために,あえて埼

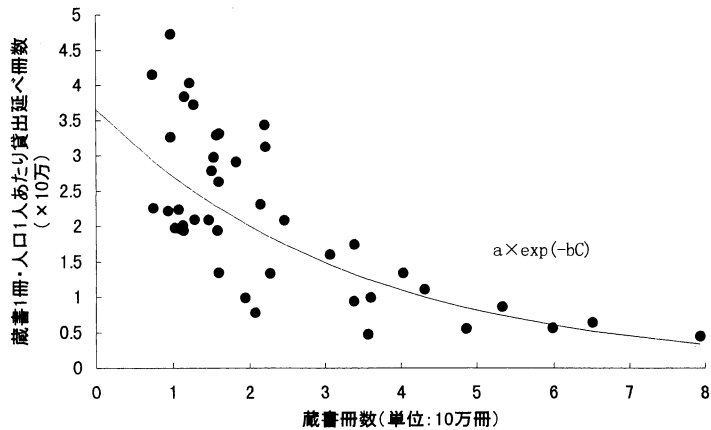


図1 埼玉県における各市の蔵書冊数と蔵書1冊・人口1人あたり貸出延べ冊数によるプロット

玉県の町村部も分析の対象からはずすことにした。

基本的な統計の概要を以下に示す。まず蔵書冊数に関しては、最も多い市が大宮市で約79万冊、最も少ない市が蓮田市で約7万冊、42市の平均で約24万冊であった。次に貸出延べ冊数については、最も多い市が浦和市で約198万冊、最も少ない市が加須市で約11万冊、平均で約49万冊であった。なお、人口 P を扱うことから、データの単位は「図書館」ではなく、「市」となる。すなわち、複数の館を持つ市の場合には、それらの蔵書冊数と貸出延べ冊数を単純合計した。

モデル(4)式は非線形であるため、このままではパラメータの推定などの分析が多少困難であることから、まず、モデル(4)式を、

$$\frac{L}{CP} = ae^{-bC} \quad (5)$$

と変形する。左辺は「蔵書1冊・人口1人あたりの貸出延べ冊数」であり、右辺には、蔵書冊数 C のみが変数として残り、形式的には負の指数関数となっている。そこで、埼玉県の42市の蔵書冊数・貸出延べ冊数・定住人口のデータを用いて、「蔵書1冊・人口1人あたりの貸出延べ冊数」を縦軸((5)式の左辺)、蔵書冊数を横軸として、各市のデータをプロットしてみる(図1参照)。

図1のグラフは比較きれいな負の指数曲線を描いている。このことから、本稿におけるモデ

ルが、このデータに対しては、近似式としてある程度妥当であると結論できよう。さらに、(5)式の両辺を対数変換すれば、 $\log [L/(CP)] = \log a - bC$ となるから、パラメータ a と b を最小2乗法で推定することができる。これを実行したところ、

$$\hat{a} = 0.0000365430\dots, \quad \hat{b} = 0.0000030004\dots$$

であった(決定係数は0.63)。この決定係数は、貸出密度(人口1人あたり貸出延べ冊数)を人口1人あたり蔵書冊数で予測した場合よりも高く(この場合の決定係数は0.50)、先行研究の結果に照らしても、説明変数が2個のモデルとしてはかなり満足のいく結果である(例えば、拙稿³⁾では、蔵書冊数、専門職比率、新聞頒布数、受入冊数、人口密度、成長力指数、図書館面積、移動図書館数、世帯人員、独身者比率の10個もの説明変数を投入し、貸出密度を予測した回帰モデルの決定係数は0.81であった)。

IV. モデルの解釈とシミュレーション

埼玉県の市立図書館の貸出延べ冊数の変動をモデル(4)式(あるいは(5)式)でかなり説明できたということは、埼玉県の市立図書館においては、いわば「規模の反有効性」とも呼ぶべき要因(係数 e^{-bC} に相当)が作用していることを意味している。これはモデルの導出過程で示したように、蔵書の規模が大きくなるにつれて、蔵書冊数の増加の効果が減衰していくことを意味している。さ

らに、蔵書冊数と人口とに高い正の相関があることを考えれば（本稿のデータでは単相関係数0.88）、蔵書冊数 C は人口規模をも表していると考えられるので、人口が大きくなればなるほど、人口の増加の効果が減っていくと捉えることもできる。すなわち、(4)式は、

$$L = aCPe^{-bP} \quad (6)$$

とすることも可能である。

特に埼玉県の一部の場合、図1が単峰形になっていないことから、「規模」が小さいほど、「規模」単位あたりの貸出量が多いという状況が一般的に成立していることになる。このことから、なるべく図書館を分割し、1館あたりのサービス対象人口を減らせば、貸出の量が増えるという結果が導かれる。これは、図書館サービスを受けるにあたっての利用者の利便性を考えれば、妥当な解釈として考えられよう。

モデル(4)式を使えば、この効果を試算することが可能である。本章では、モデルの有用性を示す1つの例として、かなり粗い仮定を置いた上で、図書館数を増やした場合の有効性を試算してみる。

図1のデータの単位は「市」であるが、図書館数の効果を調べるには、「図書館」を単位とした分析が必要になる。もちろん、『日本の図書館』には、図書館単位のデータが掲載されているが、複数の館を持つ市の場合、市全体の定住人口をそれらに割り当てることが難しいので、今回はこれを使用しない。その代わりに、粗い近似ではあるが、複数の館が設置されている市の場合には、その市の L 、 C の合計および P をその図書館数で割り、それを図書館1館のデータとしてモデルに投入することによって、擬似的に、「市」ではなく「図書館」を単位としたパラメータの推定を試みる。例えば、ある10万人の人口を持つ市に2つの図書館があって、その蔵書冊数がそれぞれ6万冊、2万冊、貸出延べ冊数が3万冊、1万冊ならば、この市には、奉仕人口が5万人($10 \div 2 = 5$)で、蔵書冊数が4万冊($(6+2) \div 2 = 4$)、貸出延べ冊数が2万冊($(3+1) \div 2 = 2$)の図書館が2館あると考える。この例から、この方法がかなり粗いものであ

ることは明らかであるが、すでに述べたように、人口規模を各館に割り当てる作業は難しく、その割当に起因する誤差が混入する可能性もある。本稿のここでの目的は、埼玉県の市部における厳密な推計ではなく、むしろモデルの有用性(使い方)の例示に比重があることから、とりあえず今回はこの方法を使って、簡単なシミュレーションを試してみることにしたい。もちろん、各図書館のサービス対象人口のより正確なデータ・推定値が利用できれば、以下の推計の精度はより高くなるわけであり、これは今後の課題としたい。

上記の方法に従ってデータを作り直し、上と同じ手順で最小2乗法によってパラメータを推定すると、

$$\hat{a} = 0.000064851\dots, \hat{b} = 0.000005880\dots$$

となる(決定係数0.54)。これは、いちおう、図書館を単位とした場合のモデル(4)式のパラメータと考えることができる。

次に、1994年の時点で市内に1つの図書館しか持っていない市(飯能市など24市)が、すべて分館を1館ずつ建設していたと仮定する。これらの市に関して、 C 、 P を2で割り、モデル(4)式に代入して、貸出延べ冊数を予測すれば、もし分館を設置していたならば、どの程度、貸出延べ冊数が増えていたかを1つのシミュレーションとして推計できる。 C を単純に2で割るということは、「市の全体としての蔵書冊数は不変のまま、単に図書館数が多かったならどうなっていたか」という問題を考えることを意味している(もし実際に図書館数を増やせば、蔵書冊数 C も増加するはずだが、ここでの推計はそのような状況を考えるものではないことに注意)。

この計算を試みた結果、これら24市における1市あたりの平均貸出延べ冊数は、

実際: 297,824.0冊

シミュレーションの結果: 314,996.0冊

であり、約5.8%の増加となることが試算された(具体的には、モデルに従って個別に各市の予測値を算出し、それを合計して24で割ることにより、平均貸出延べ冊数を算出している)。これがモデル(4)式より導かれる「アクセスポイントとし

ての図書館数の増加」の効果である。なお、すでに述べたように、実際に1つの分館を増設する場合には、市全体の蔵書冊数はそれなりに増加するであろうから、その時点までの蔵書冊数を単純に2で割った数値を用いた場合の貸出延べ冊数の予測値はかなりの過小評価になっていると考えられる。そこで、この推計を実際の図書館政策に役立てるには、新図書館建設に伴う付随的な蔵書冊数の増加を実情に応じて設定した上でモデル(4)式を利用し、費用対効果分析等を試みるが必要になる。

もちろん、単館の市ばかりでなく、すでに複数の分館を設置している市においても、利用者の利便性の向上の観点からは、さらに分館を設置することが望ましい。このような場合の試算にも、本稿のモデルは当然利用可能である。

V. おわりに

本稿では、公共図書館の貸出に関する独自の演繹的モデルを作成し、埼玉県公共図書館について分析した。このモデルは説明変数として蔵書冊数とサービス対象人口(定住人口)とを含み、それらの規模に対応した減衰要因を組み込んでいる点に特徴がある。実際のデータを使った分析の結果、埼玉県の市部においては、このモデルがある程度適合し(図1参照)、規模の反有効性と呼ぶべき効果が埼玉県の市部の公立図書館において作用していることが明らかとなった。図1で示したような分布の形状は他の地域でも成り立つ可能性があるものの、本稿で提示したモデルが埼玉県の町村部あるいは他の地域でも近似式として妥当であるかどうかは今後の研究課題である。

モデルに関しては、さらにパラメータを導入して一般化することも考えられる。例えば、パラメータ d, k を導入して、 $L = aC^d P^k e^{-bc}$ のようにすれば、説明力は確実に増すであろう。しかし、この場合には、もはや古典的最低2乗法を適用することはできないので、パラメータの推定の方法の問題を解決しなければならない。さらに、第II章で指摘したように、蔵書冊数・定住人口以外の説明変数をモデルに組み込むことも将来的な課題

である。この場合には、モデルの演繹的導出過程をかなり修正する必要が出てくる可能性がある。

また、本稿ではモデル(4)式(あるいは(5)式)の有用性を示すために、粗い仮定に基づくシミュレーションを試み、アクセスポイントとしての図書館数の増加の効果について試算した。その結果、1994年時点で単館のみを設置していた埼玉県の市がそれぞれもし2館の図書館を持っていたならば、貸出延べ冊数は約5.8%増加していたはずであると推定された。第IV章で再三指摘したように、この推計はいくつかの非現実的な仮定に依拠しているため、信頼性の点では低く、かなりの過小評価となっている可能性がある。そこで、もし将来的に本稿のモデルが認められた場合、このシミュレーションの方法についても、精緻化を進めていく必要があると考えられる。

謝 辞

本稿は、平成7年度および8年度駿河台大学共同研究助成費「埼玉県西部山麓地域の文化構造の研究」(代表: 広瀬順皓文化情報学部教授)における研究成果の一部をまとめたものである。この助成費を利用して、(社)日本図書館協会から『日本の図書館』の機械可読データの提供を受けた。また、本稿は、三田図書館・情報学会1997年度研究大会(1997年10月18日、慶應義塾大学三田キャンパス)における研究発表「公共図書館の貸出に関する演繹的モデル」を短報としてまとめたものである。この発表に対して、当日、多くの方々からご批判・ご示唆をいただきました。特に、国立環境研究所の大井鉦先生には、貴重なご意見をいただくとともに、関連研究をお教えいただきました。ここに記して感謝いたします。

注・引用文献

- 1) 1970年代から80年代かけて、多くの統計的研究がなされた。例えば、次の文献などを参照。Zweizig, D. and Dervin, B. Public library use, user, uses: advances in knowledge of the characteristics and needs of the adult clientele of American public libraries. *Advances in Librarianship*. Vol. 7, p. 231-255 (1977)
- 2) 糸賀雅児。"公共図書館の活動指標と図書館内の

- 要因の分析”. 図書館学会年報. Vol. 28, p. 13-28 (1982)
- 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. No. 24, p. 45-55 (1986)
 - 4) 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. No. 28, p. 121-143 (1990)
 - 5) 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. No. 29, p. 161-168 (1991)
 - 6) 例えば次の文献など. 近藤延幸・前田良人. 市町村立図書館の経営分析に関する考察と鳥取県立図書館の現状. 鳥取, 鳥取県立鳥取図書館, 1987. 133 p.
 - 7) 田村 肇. “都道府県立図書館の活動指標と利用要因の間の因果関係に関する実証的分析”. 情報の科学と技術. Vol. 46, No. 4, p. 227-238 (1996)
 - 8) 図書館計画・建築の領域では, 利用者の住居と図書館までの距離や, 最寄り駅との位置関係が図書館利用の重要な要因となるという研究結果がある. したがって, 自治体の規模が大きくなれば, それだけ, 図書館と利用者の居住地が離れ, その地域の住民の利用率は減少することになる. 住居までの距離に関する最近の研究例としては以下のものがある. 武田 直・尾崎幸子・島崎敏一. 公共図書館の利用と設置状況. 土木学会第 52 回学術講演会, 1997 年 9 月 (講演概要集第 4 部 p. 484~485)
 - 9) 日本図書館協会編. 日本の図書館 1994. 東京, 日本図書館協会, 1994. 404 p.