Library and Information Science No. 28 1990

蔵書構築のための計量的手法の基礎としての貸出頻度分布モデルの有用性

On the Usefulness of a Model of Frequency Distribution of Book Circulation as a Basis for Quantitative Approaches to Collection Development

岸 田 和 明
Kazuaki Kishida

Résumé

Recently some models for frequency distribution of book circulation have been developed. These library circulation models have the potential capability of being a basis for quantitative approaches to collection development. Namely, these models describe the use of books more subtlely than other models, so we can obtain an improvement to quantitative methods for collection development by applying a model of the frequency distribution of book circulation to it.

In this paper, firstly, it is shown that the model for frequency distribution of book circulation contains rich information on the use of books. By discussing some mathematical forms of obsolescence, it is shown that Brooke's utility, Morse's circulation model and Bradford's law are represented by the model for frequency distribution of book circulation. Particulaly Burrell's non-stationary Poisson process model is used as an example of a model for frequency distribution of book circulation. Secondly, by applying this model to Morse's queuing theory method for multiple copies and Trueswell's method for the weeding of books, these methods are extended. This implies the applicability of the model for frequency distribution of book circulation to quantitative methods for collection development.

岸田和明：図書館情報大学助手，茨城県つくば市春日1-2
Kazuaki Kishida: Assistant, University of Library and Information Science, Tsukuba-shi, Ibaraki-ken, Japan

I. はじめに

近年の出版量の増大，物価の上昇などの諸要因により現在，各図書館はその予算やスペースに著しい制約を受けている。もはや各館が単独で世の中のすべての資料を収集・保存することは不可能であり，たとえサービス対象とする主題分野を狭い範囲に限定したとしても，その利用者の要求すべてを満足させるだけの資料を単独で維持することは非常に困難となっている。

このような状況に対処するためには，受け入れる資料の選択，保存書庫への資料の別置，不要な資料の廃棄などを行い，制限されたスペース・予算の中で利用者を最大限に満足させるように努力する必要がある。これらの作業は，蔵書構築と呼ばれる。蔵書構築 (collection development) という概念の捉え方は現在の図書館・情報学分野においてはまだ十分に安定したものではないが，本論文では，上に述べたような状況に対応して，従来の図書選択 (book selection) の概念が拡張され，廃棄や別置，複本購入などの要素が加えられたものとして蔵書構築を捉える。つまり，新たな資料の受け入れだけに注目するのではなく，その結果として形成された蔵書自体の後の時点での適切性をも問題とするのである。利用者への資料の提供という図書館の目的を考えれば，この意味での蔵書構築は図書館経営にとって非常に重要な部分であると言える。

この蔵書構築において，最終的な決定を行うのは，図書館員である。そこで，図書館員にはその決定を行うための能力・技術が要求される。しかし，時として個々の図書館員の判断力だけでは決定が困難となる場合や，図書館員の主観だけでは限界のある場合が出てくる。例えば，開架書庫を縮小するために別置する資料を選択する場合や，新たに追加して購入する雑誌を決める場合などは，図書館員の主観だけで選択を行うことは難しい。このような場合，その分野の専門家の意見を求めるという方法がある。

しかし他の有力な方法として，資料の利用回数や被引用回数のような量的な尺度を用いることが考えられる。実際，先に掲げた例では，利用回数や被引用回数が選択のための有力な情報となるであろう。さらに，資料選択の際の各分野への予算の配分問題や，書庫スペースの耐用年数の算出などの問題になれば，図書館員や専門家の主観的な判断よりも，むしろ統計学やオペレーションズ・リサーチなどの計量的な手法に全面的に依拠したほうが，効果的かつ効率的であると考えられる。

このように，計量的な尺度や手法は図書館員が行う蔵書構築に関する種々の意思決定の補助として非常に有用である。そこで，現在までに統計学やオペレーションズ・リサーチにおける手法が蔵書構築に導入されたり，あるいは新たに計量的な手法が図書館・情報学分野で開発されたりしてきた。例えば，Morse[1] の複本購入のための待ち行列モデルなどは，オペレーションズ・リサーチの分野で開発された理論を蔵書の問題に適用した典型例である。

しかし，それらの計量的手法は十分であるとは言えない。例えば，先の Morse の待ち行列モデルは，ある年次 t の貸出回数および貸出期間の数値から，複本購入の効果等を推定するモデルであるが，実際には obsolescence (経年的な利用や引用の減少傾向) の効果があるのでモデルから推定された結果は $t+1$ 年の状況に適用することができない。つまり，obsolescence という図書館・情報学でよく知られている現象をモデルに組み込んでいないために，モデルの有効性が半減してしまっているのである[2]。

これは，ほんの一例にすぎないが，この問題の背景に

は，文献の利用あるいは引用を正確に記述する基礎的な数学モデルがないという状況があると考えられる。例えば上の例では，文献の貸出の時系列的な変化を記述するモデルを待ち行列モデルに組み込めば，上述した問題点は解決される。図書館が利用されることを目的としている以上，文献が利用される回数や引用される回数は，図書館におけるその文献の有用性を測る尺度として有効であると考えられる。そこで，それらを正確に記述する基礎的なモデルがあれば非常に便利であるし，逆にそのようなモデルがなければ，実際に有用な計量的手法を開発することは不可能であるとも言えよう。

この文献の利用あるいは引用を記述するモデルの候補としては，ブラッドフォードの法則を表現する関数やobsolescenceを表現する関数が考えられる。しかし，これらは十分ではない。例えば，ブラッドフォードの法則を表現するいくつかの関数は，文献の利用回数や被引用回数の分布を記述するが，その経年変化は扱うことができない[3]。一方，obsolescenceを表現する指数関数は，総利用回数や平均利用回数などの，文献集合のマクロ的な特性に関しての経年変化を記述する場合に，その文献集合中の個々の文献を区別して扱うことはできない。そこで，これらの関数は蔵書構築のための計量的手法の基礎となるモデルとはなり得ない。

しかし，一方，最近研究の進められている「貸出頻度分布モデル」が，ある程度その条件を満たしている可能性がある。「貸出頻度分布」とは，貸出可能な図書の貸出回数に関する度数分布を指す。そして，これを確率分布を用いてモデル化したものを本論では「貸出頻度分布モデル」と呼ぶ（確率変数が貸出回数，確率が図書の冊数の全体に対する割合に相当する）。この貸出頻度分布は単純な度数分布に過ぎないが，度数分布は事象を計量的に記述する場合の基本であり，それだけに蔵書構築に関する計量的手法の基礎としての役割を果たしうる可能性がある。

そこで本論文では，蔵書構築のための計量的手法における基礎的なモデルとしての，貸出頻度分布モデルの有用性について考察を進めることにする。このことが明らかになれば，蔵書構築のための計量的手法やモデルの開発にとって重要な示唆が得られることになるだろう。

しかし，蔵書構築のための計量的手法における基礎的なモデルであるための必要条件を明示することは非常に難しい。そこで本論文では，いくつかの事例から次の二点を確認することによって，貸出頻度分布モデルの基礎的モデルとしての有用性を明らかにしたい。その二つの確認事項とは，

①現在までに蔵書構築にとって有用であると考えられているいくつかの計量的な手法あるいは尺度，モデルを貸出頻度分布モデルを用いて表現できること，あるいは，貸出頻度分布モデルがそれらの手法や尺度，モデルと同じような機能・役割を果たすことができること，

②貸出頻度分布モデルが，現在までに開発されている何らかの手法・モデル内に，利用を表現する関数として組み込まれた結果，その手法・モデルが精緻化あるいは拡張されること，

である。基礎的なモデルは，利用の状況に関して豊富な情報を持っている必要がある。そこで，利用の状況に関して，現在までに用いられている尺度やモデルと同等あるいはそれ以上のことを表現できなければならない。これが上の①にあたる。一方，いくら豊富な情報を持っていようと，その情報を実際に幅広く応用できなければ意味がない。そこで基礎的なモデルは他の蔵書構築のための手法やモデルにも応用可能でなければならない。このことが②に相当する。

実際，①に関しては，第Ⅲ章において，蔵書回転率のような指標やブラッドフォードの法則の数式表現などが貸出頻度分布モデルによって表されることを示す。そしてその結果として，それらの指標・モデルが貸出頻度分布モデルのひとつの特殊形にすぎないことが明らかになる。また②に関しては第Ⅳ章において，上で述べたMorseのモデルに貸出頻度分布モデルを組み込むことによって，そのいくつかの欠点が修正されることを中心に論じる。

しかし，その反面，貸出頻度分布は貸出可能な図書の貸出に限定されたものであり，雑誌や引用回数・館内利用が扱えないという欠点がある。これは確かに大きな問題ではあるが，図書の貸出は現在のところ，図書館経営における量的な指標として大きな比重を占めているし，また，貸出頻度分布モデルを「館内利用頻度分布」「引用頻度分布」などに拡張する可能性も残っている。そこで，貸出頻度分布を用いることによって対象が図書の貸出に限定されることは，本研究の試みの価値をそれほど大きく損なうものとは考えられない。ただし，現時点ではその拡張のことは考慮しないので，本論文で扱う図書館資料は，貸出可能な図書のみである。

II. 貸出頻度分布モデル

A. 貸出頻度分布モデルの条件

第I章で述べたように，貸出頻度分布とは貸出回数に関する度数分布のことである。図示すれば，横軸に貸出回数，縦軸にその冊数（度数）をプロットすることになる。一方，貸出頻度分布モデルは，この度数分布を確率分布を用いてモデル化したものである。

本論文の目的は，この貸出頻度分布モデルが他の計量的尺度や手法，モデルを表現すること，あるいはそれらと同等以上の機能を果たすことを示すことにある。この目的のためには，貸出頻度分布モデルが単なる度数分布ではなく，貸出頻度分布の経年的な変化を記述することが可能なものでなければならない（その理由は第III章で明らかになる）。そこで，本論文で用いる貸出頻度分布モデルは，x と t の二つの変数を持つ関数となる。ここで，変数 x は貸出回数であり，t は時間である。したがって，貸出頻度分布モデルは $f(x, t)$ と表記される。

貸出頻度分布の確率分布によるモデル化にあたっては時間 t をある時間間隔に限定した特別な確率変数を考えておくと便利である。それは，

①X_t: 初期時点から時点 t までの貸出回数
②Y_t: 時点 $t-1$ から時点 t までの貸出回数
③$Z(t, h)$: 時点 t から時点 $t+h$ までの貸出回数

の三種類である。これらの有用性は，時間 t の単位を年で考えるとわかりやすい。例えば，Y_t は第 t 年次の一年間の貸出頻度分布に対応する。また，X_t は最初の時点から第 t 年次までの貸出頻度分布，同様に $Z(t, h)$ は第 $t+1$ 年次から第 $t+h$ 年次までの貸出頻度分布に対応する。本論文ではこれ以降，便宜的に，このように時間 t の単位を年でとり，$f(x, t)=f(Y_t)$ を「各年次分布」，$f(x, t)=f(X_t)$ を「累積年分布」と呼ぶことにする。さらに，$Z(t, \mathrm{h})$ からは，条件付き確率分布を使用して，

$$f(Z(t_2, h_2)|Z(t_1, h_1))$$

を考えることができる。これは，t_1 年以降 h_1 年間に $Z(t_1, h_1)$ 回貸出された図書が，さらに t_2 年以降 h_2 年間に $Z(t_2, h_2)$ 回貸出される確率である。

第III章での議論で対象となる $f(x, t)$ は，本質的には上の $f(X_t)$, $f(Y_t)$, $f(Z(t, h))$ あるいは $f(Z(t_2, h_2)|Z(t_1, h_1))$ を表現できるものならば，どのようなものでもよい。しかし，実際にそのような条件を満たす貸出頻度分布モデル $f(x, t)$ が具体的に設定できることを確認しなければ，第III章での議論は，単なる空論に終わってしまう。そこで，図書館・情報学分野において上の条件を満たす $f(x, t)$ についての具体的なモデルが提案されていることと，そのモデルが妥当であることとを確認しておく必要がある。

したがって，本章では次節以降，上の条件を満たす貸出頻度分布モデル $f(x, t)$ として，負の二項分布に基づくモデルが実際に存在し，それがある程度妥当であることを見ることにする。そのために，まず B 節で負の二項分布が貸出頻度分布を十分に記述することを先行研究の結果から議論し，次に，C 節でその負の二項分布に時間変数 t を組み込んだ，経年変化を記述する貸出頻度分布モデルについて論じる。このような具体的なモデルを論じておくことは，第III章での議論をより具体的なイメージの下に進めていくことにも役立つであろう。

B. 貸出頻度分布モデルとしての負の二項分布の妥当性

歴史的には，Morse[1] がまず貸出頻度分布として幾何分布を用いた。しかし，貸出頻度分布を幾何分布で記述しようとした Burrell[4] および Burrell and Cane[5] に対して，負の二項分布の方が優れているという批判がなされた[6]。このような状況を背景として，Bagust[7] が負の二項分布の妥当性を実際の公共図書館の貸出データから検証し，さらに Burrell[8] が負の二項分布に関するより優れた考察を行うにしたがい，負の二項分布の優位性が次第に認められるようになった。

現在，負の二項分布以外に貸出頻度分布モデルの候補となる確率分布としては，ベータ二項分布と逆ガウス-ポアソン分布とがある。ベータ二項分布は，Gelman and Sichel[9] により貸出頻度分布を記述する確率分布として提案された。一方，逆ガウス-ポアソン分布は，Sichel[10] がジップの法則やロトカの法則よりも正確に単語の出現回数や研究者の論文発表数を記述する確率分布として発表したものであり，最近その有用性が認められつつある[11]。

しかし，負の二項分布とこれら二つの確率分布との比較を行った岸田ほか[12] および岸田[13] の結果によれば，貸出頻度分布としては負の二項分布が優れていると結論することができる。前者の研究は，幾何分布，対数級数分布，ネイマンのA型伝播性分布，ベータ二項分布と負の二項分布とを実際のデータへの適合度で比較し，後者の研究は，ツェータ分布，ユール分布，逆ガウス-ポア

ソン分布などと負の二項分布とを比較しているが，どちらの場合も負の二項分布の方が適合度が高かった[14)]。

すべての図書館で負の二項分布が適用できるという一般的な結果はまだ得られていないものの，以上挙げた研究結果からは，かなりの図書館で負の二項分布を適用することができると考えられる。

C. 非定常ポアソン過程型貸出頻度分布モデル

負の二項分布が複合確率分布

$$f(x)=\int h(x;r)g(r)dr \qquad (1)$$

として導出されることは一般に知られているが，Burrell[4)] は，この複合確率分布の考え方を貸出頻度分布を説明する理論として応用した。まず，各図書のある一定期間の貸出回数がポアソン分布

$$h(x;r)=e^{-r}r^{x}/x\,! \qquad (2)$$

で表現されると仮定する。次に (2) 式のパラメーター r が各図書によって異なり，この r 自体が蔵書全体である確率分布 $g(r)$ にしたがって分布していると仮定する。(2) 式のパラメーター r は一冊の図書が一定期間に貸出される回数の平均（期待値）と解釈できるので，Burrell は r を「望ましさ (desirability)」と呼んだ。つまり彼は，一冊の図書の貸出はその図書の以前の貸出に関係なくランダムに生起するが，その貸出の平均回数が各図書によって異なるという構造から貸出頻度分布が成り立っていると考えた。この構造は (1) 式で表現される複合確率分布に対応する。そこで，$g(r)$ として何らかの確率分布を具体的に設定すれば，(2) 式を用いて (1) 式を計算し，貸出頻度分布モデル $f(x)$ を求めることができる。例えば前節で述べた負の二項分布は $g(r)$ をガンマ分布とすれば，(1) 式より導出される。

この複合確率分布によるモデル化が重要なのは，(1) 式における $h(x;r)$ を $h(x,t;r)$ とすることによって，貸出頻度分布の時間的な変化を記述するモデルの構築が可能になるからである。Burrell[8)] は，(2) 式の代わりに

$$h(x,t;r)=e^{-M(r,t)}M(r,t)^{x}/x\,! \qquad (3)$$

を用いた。ただし，

$$M(r,t)=\int_0^t re^{-as}ds \qquad (4)$$

とおく[8)]。$a>0$ ならば明らかに，この関数 $M(r,t)$ は時間 t の経過によって減少する。先に述べたように，ポアソン分布のパラメーターは，この場合平均貸出回数として捉えることができるので，結局 (4) 式は obsolescence を表現していることになる。

そこで，(3) 式を (2) 式の代わりに (1) 式に代入して計算すると，貸出頻度分布の時間変化を記述するモデル $f(x,t)$ を得ることができる。これを「非定常ポアソン過程型貸出頻度分布モデル」と称する。なぜなら，(3) 式は時間変化するパラメーターを持つポアソン分布であるので，この式で表現される確率変数は「非定常ポアソン過程」に従っていると見なすことができるからである。特に，先に述べたように $g(r)$ としてガンマ分布を仮定すると，貸出頻度分布を記述する負の二項分布が得られ，非定常ポアソン過程型貸出頻度分布モデルは，その負の二項分布のパラメーターが時間変化するかたちとなる。

このようにして得られた非定常ポアソン過程型貸出頻度分布モデル $f(x,t)$ は，ある図書が初期時点から時点 t までに x 回貸出される確率を与える。すなわち，本章 A 節で述べた $f(X_t)$ である。具体的には，

$$f(X_t=m)=\binom{m+k-1}{m}p(t)^k[1-p(t)]^m,$$
$$m=0,1,2,\ldots \qquad (5)$$
$$p(t)=[1+\frac{a}{b}(1-e^{-at})]^{-1}$$

である[8)]。ここで，k と b はガンマ分布 $g(r)$ のパラメーター，a は (4) 式のパラメーター (obsolescence の係数) である。これが，経年変化を記述する貸出頻度分布モデル $f(x,t)$ の具体的なかたちのひとつである。

さらに，この非定常ポアソン過程型貸出頻度分布モデルにおいては，

$$f(Y_t)=f(X_t-X_{t-1}) \qquad (6)$$

$$f(Z(t,h))=f(X_{t+h}-X_t) \qquad (7)$$

と考えることにより，$f(Y_t)$, $f(Z(t,h))$ も $f(X_t)$ と同様に導出することができる。

まず $f(Y_t)$ の場合は，(6) 式に従って，(4) 式の積分区間を $(t-1,t]$ に変えればよく，その結果モデルは，

$$f(Y_t=m)=\binom{m+k-1}{m}p_t^k[1-p_t]^m,$$
$$m=0,1,2,\ldots \qquad (8)$$
$$p_t=[1+\frac{b}{a}e^{-a(t-1)}(1-e^{-a})]^{-1}$$

となる[8)]。同様に，$f(Z(t,h))$ の場合は積分区間を (7) 式

に従って，$(t, t+h]$ に変え，

$$f(Z(t,h)=m)=\binom{m+k-1}{m}p(t,h)^k[1-p(t,h)]^m, \quad m=0,1,2,3,\ldots \quad (9)$$

$$p(t,h)=[1+\frac{b}{a}e^{-at}(1-e^{-ah})]^{-1}$$

と計算される[15)16)]。

一方，条件付き確率分布 $f(Z(t_2, h_2)|Z(t_1, h_1))$ に関しては，条件付き確率分布の理論を用いて，

$$f(Z(t_2,h_2)=m|Z(t_1,h_1)=n)$$
$$=\binom{m+n+k-1}{m}\pi(t_1,t_2,h_1,h_2)^k[1-\pi(t_1,t_2,h_1,h_2)]^m, \quad m=0,1,2,3,\ldots \quad (10)$$

$$\pi(t_1,t_2,h_1,h_2)=\frac{bG(t_1,h_1)+1}{bG(t_1,h_1)+bG(t_2,h_2)+1}$$

$$G(t,h)=\frac{1}{a}e^{-at}(1-e^{-ah})$$

となる[15)]。

これで，Burrell[8)15)16)]の非定常ポアソン過程型の貸出頻度分布モデルが，本章 A 節で述べた条件件を満たしていることがわかった。

D. 貸出頻度分布モデルに関する注意点

以上のように，貸出頻度分布は確率分布を用いてモデル化される。しかし，この貸出頻度分布モデルに関してはいくつか注意点がある。

まず第一に，経年変化を予測する場合，対象とする図書の集合は，時点 $t=0$ で固定する必要がある。つまり，ある図書集合に関する $f(x,t)$ のパラメターを求めた場合，後から図書を加えることができない。図書を加える場合は，その加えた時点を $t=0$ として，再び $f(x,t)$ のパラメターを求め直す必要がある。このために，図書館の蔵書全体を $f(x,t)$ の対象とするのには問題があり，蔵書を受け入れ年別に分割して考えた方がよい。つまり第 s 年に受け入れられた図書に関する貸出頻度分布モデルを $f_s(x,t)$ とするのである。出版年と受け入れ年を等しいと考えることができるならば s は出版年でもよい。そして蔵書全体の貸出頻度分布は，$f_s(x,t)$ が集まったものと見なす。

ここで，指数 s と変数 t の取り方にも注意する必要がある。t は受け入れ（出版）の時点 s を 0 として増加する。したがって，例えば，1970年と1980年に受け入れられた図書の1990年におけるそれぞれの貸出頻度分布を考える場合，それぞれの変数 t には 10 の差があることになる。ある時点 s のみの貸出頻度分布を考える場合には，この問題が起こらないが，複数の貸出頻度分布モデルを同時に扱う場合で，t を変数とする（固定しない）ときには，t の値の取り方について気をつける必要がある。

第二に，貸出頻度分布モデルの解釈について注意する必要がある。つまり，本来貸出頻度分布は単なる度数分布であったが，それを確率分布を用いてモデル化したために，結果として作成された貸出頻度分布モデルはさらに確率的に解釈することが可能になったのである。例えば，(5)式の $f(X_t=m)$ は，$[0,1]$ の実数で与えられる。そうすると，(5)式は，対象とする蔵書全体から一冊の図書を無作為に選んだとき，その図書の貸出回数が m 回である確率を与えていると考えられる。一方，(5)式の $f(X_t)$ に蔵書数をかければ，度数分布が得られるから，(5)式は単に度数分布を表現しているとも見なすことができる。つまり，貸出頻度分布モデル $f(x,t)$ は，

①一冊の図書を選び出したとき，それが x である確率を与えるもの

②度数分布を単に図書総数で割って正規化したもの

という二種類の解釈が可能なのである。本論文では，この二つの解釈を場合に応じて使い分けることにする。

III. 蔵書構築のための計量的手法と貸出頻度分布モデル

本章では，現在までに蔵書構築にとって有用であると考えられているいくつかの計量的手法あるいは尺度，モデルが，貸出頻度分布モデルによって表現されることを見る。これによって，貸出頻度分布モデルが貸出に関する豊富な情報を持っていることが示される。

A. obsolescence と貸出頻度分布モデル

図書館において，古くなった資料が利用されなくなるという事実が一般に知られている。この現象は「obsolescence」あるいは「aging」と呼ばれ，多くの研究が行われている。これは，古くなり利用されなくなった資料を廃棄あるいは別置の候補として考えることができるためである。

この obsolescence の数学的なモデル化を最初に試みたのは1944年の Gosnell[17)]であり，彼はすでに obsolescence を指数関数を用いて表現していた。しかし，

さらに obsolescence の研究に重要な進展をもたらしたのは，1960年の Burton and Kebler[18] であり，彼らは科学文献の obsolescence と放射性物質の崩壊との類似性から「半減期 (half-life)」という指標を提案した。

以来，数多くの obsolescence に関する研究が行われており，現在では obsolescence を表現するには一般に

$$y(t)=Ke^{-at} \tag{11}$$

という式が用いられる。$y(t)$ は t 年後の利用回数あるいは被引用回数を表し，K と a はパラメターである。

ただし，(11)式をデータに適用する際，二通りの適用の方法があることに注意しなければならない。第一は何らかの文献集合に対しての利用回数・被引用回数に(11)式を適用する方法，第二はあるひとつの資料のみに限定し，その利用や被引用回数に対し(11)式を適用する方法である。例えば，ある年に出版された図書の貸出回数を問題とする場合は前者にあたり，ある特定の雑誌に対する引用を扱う場合は後者に該当する。図書の貸出の場合は，前者の方法をとることがほとんどである。そこで本節では，図書を集合として扱った場合の obsolescence が貸出頻度分布モデルで表現されることを見る。

しかし，obsolescence の捉え方にはいくつかの種類があるので，それらを区別しておく必要がある。まず第一は「共時的 (synchronous) な obsolescence」と「通時的 (diachronous) な obsolecence」との区別である[19]。通時 obsolescence はある文献集合の利用や引用の減少を経年的に見たものであり，一方，共時 obsolescence はある単位時間（通常一年間）において利用または引用された文献についての刊行年の分布を対象とする。第二は文献数で補正するか否かの区別である。これはつまり，利用可能あるいは被引用可能な文献の数で，利用回数あるいは被引用回数を補正するかどうかの問題である。この問題は，文献数が経年的に増加するために生じたものである。

さて以上の点に注意して，図書を集合として扱った場合の obsolescence が貸出頻度分布モデルで表現されることを示す。このために，ここでは $f(x,t)$ として各年次分布 $f(Y_t)$ を用いる。ただし，特に出版（受け入れ）年を示す指数 s を用いて，$f_s(Y_t)$ とする（前章 D 節参照）。

まず，共時 obsolescence を考える。先に述べたように共時 obsolescence は出版年の分布であるから，s を変数として，$\Phi(s)$ と表す。一方，t は，$\Phi(s)$ がある時点に固定されるように適当に設定する。例えば，1990年における $\Phi(s)$ を問題とする場合，1980年に出版された図書の $f(x,t)$ の場合は $t=10$，1985年に出版された図書の場合は $t=5$ と自動的に設定される。さて，この点に注意すれば，文献数で修正しない場合，第 t 年次の共時 obsolescence $\Phi(s)$ は，

$$\Phi(s)=\sum_{Y_t=0}^{L} C_s Y_t f_s(Y_t)\,, \quad (t\text{ は固定}) \tag{12}$$

と表すことができる。ここで，C_s は第 s 年に出版された（受け入れられた）図書についての蔵書数，また L は変数 Y_t の最大値である。(12)式は s 年に出版された図書が第 t 年において貸出された回数の合計を意味している（t は上で述べたように設定）。そして s が変数だから，(12)式はある一時点の貸出における出版年の分布，すなわち共時 obsolescence を表現することになる。一方，文献の修正を考える場合の obsolescence $\Psi(s)$ は，$\Psi(s)=\Phi(s)/C_s$ であるから，これを(13)式に用いると

$$\Psi(s)=\sum_{Y_t=0}^{L} Y_t f_s(Y_t)\,, \quad (t\text{ は固定}) \tag{13}$$

となる。この式から，文献の修正を考える場合の共時 obsolescence は，t を固定した場合の，$f_s(Y_t)$ の期待値の s の増加に伴う変化に等しいことがわかる。

次に，通時 obsolescence を考える。文献数で修正しない場合，第 s 年に出版された図書の第 t 年次の通時 obsolescence $\Phi^*(t)$ は，s を特定の値に決めて，

$$\Phi^*(t)=\sum_{Y_t=0}^{L} C_s Y_t f_s(Y_t)\,, \quad (s\text{ は固定}) \tag{14}$$

である。この式は s 年に出版された図書が t 年において貸出された回数の合計を意味する。ここで，t が変数だから，(12)式とは対照的に，(14)式は通時 obsolescence を表現している。同様に，文献で修正する場合は，

$$\Psi^*(t)=\sum_{Y_t=0}^{L} Y_t f_s(Y_t)\,, \quad (s\text{ は固定}) \tag{15}$$

となる。やはり，$\Psi^*(t)$ は s を固定した場合の $f_s(Y_t)$ の期待値に等しい。

したがって，共時 obsolescence と通時 obsolescence は(12)式と(14)式とを見比べればわかるように，式のかたちは等しく，s を固定するか，t を固定するかだけの相違であることがわかる。そこで，両方を変数と考えると，

$$\Phi(s,t)=\sum_{Y_t=0}^{L} C_s Y_t f_s(Y_t)$$

$$\Psi(s,t)=\sum_{Y_t=0}^{L} Y_t f_s(Y_t)$$

とおくことができる。

(12)～(15)式は，図書の貸出に関して，従来用いられてきた obsolescence の尺度を完全に表現している。これで，貸出頻度分布モデルが obsolescence の尺度を表現することが示された。

次にこのことを具体的に，非定常ポアソン過程型貸出頻度分布モデルで考える。(8)式は基本的には負の二項分布であるから，その期待値 $E(Y_t)$ は，

$$E(Y_t)=k(1-pt)/pt$$
$$=\frac{kb}{a}(1-e^{-a})e^{-a(t-1)}$$

となる[8]。そこで，$K'=kb(1-e^{-a})/a$ とおけば，

$$\Psi^*(t)=K'e^{-a(t-1)} \qquad (16)$$

が得られる。これは，既に述べた obsolescence を表現するのによく用いられる負の指数関数である（(11)式参照)。この式が得られたことは，実は(4)式の被積分関数を e^{-at} とおいたことに起因しているが，貸出頻度分布モデルから(14)式を用いて，(16)式のようなよく知られた obsolescence の式が直接得られたことは重要な結果であると言える。

B. Brookes の効用関数と貸出頻度分布モデル

一方，Brookes[20]は，obsolescence が(11)式のような指数関数で表されることを生かして，「効用 (utility)」を提案した。(11)式において $e^{-a}=d$ とおく (Brookes は d を老化係数 (aging factor) と呼んだ)。この d を用いると，出版後その資料に対してなされる利用あるいは引用の総数 $U(0)$ は，

$$U(0)=K(1+d+d^2+d^3+\cdots+d^t+\cdots)$$
$$=K/(1-d) \qquad (17)$$

と求めることができる。同様に，出版後 t 年以降，その資料が利用・引用される総数 $U(t)$ は，

$$U(t)=K(d^t+d^{t+1}+d^{t+2}+\cdots)$$
$$=Kd^t(1+d+d^2+d^3+\cdots+d^t+\cdots)$$
$$=d^tU(0) \qquad (18)$$

となる。この U を Brookes は「効用」と呼んだ。通時的データに対してこの U を用いれば，その資料の利用・引用の将来予測が可能になり，U の小さい資料を廃棄・別置の候補とすることができる。

同様の計算を貸出頻度分布モデルから導かれた(16)式に対しても行うことができる。しかし，貸出頻度分布モデルを用いれば，さらに精密な計算を行うことが可能である。

まず(17)式に対しては，累積年分布 $f(X_t)$ において $t\to\infty$ とすれば，その図書集合についての，時点 $t=0$ から時点 $t=\infty$ までのすべての貸出を表現する貸出頻度分布を $f(X_\infty)$ として予測することができる。そして，その平均 $E(X_\infty)$ によって，一冊あたりの総貸出回数を求めることができる。また，$E(X_\infty)$ に図書数をかければ，その図書集合に対する総貸出回数も予測できる。

一方，(18)式に対しては，$f(Z(t,h))$ を用いて同様の計算を行えばよい。すなわち，$Z(t,h)$ は時点 t から時点 $t+h$ までの貸出回数であるから，$h\to\infty$ とすれば，時点 t 以降（すなわち $t+1$ 年次以降）の貸出頻度分布を予測することができる。

このことを具体的に，非定常ポアソン過程型貸出頻度分布モデルで示す。まず，$f(X_t)$ である(5)式において，$t\to\infty$ とすると，パラメター $p(t)$ は，

$$p(\infty)=[1+b/a]^{-1} \qquad (19)$$

となる。この式をパラメターとする貸出頻度分布モデルが，$f(X_\infty)$ である。次にこの $f(X_\infty)$ の平均を計算すれば，

$$E(X_\infty)=k[1-p(\infty)]/p(\infty)$$
$$=kb/a \qquad (20)$$

となる。これらが(17)式の「効用」に対応する。

一方，$f(Z(t,h))$ である(10)式で $h\to\infty$ とすると，

$$p(t,\infty)=[1+\frac{b}{a}e^{-at}]^{-1} \qquad (21)$$

であり，上と同様に平均は，

$$E(Z(t,\infty))=(kb/a)e^{-at} \qquad (22)$$

となる。これらは(18)式に対応する。これで貸出頻度分布モデルから，「効用」と同様の数値を計算できることが示せた。

しかし，(19)～(22)式を実際の廃棄や別置に応用しようとすると，不都合が生じる。これは，図書を単体としてではなく，ひとつの集合体として扱い，その集合体の

中では個々の図書を区別しなかったためである。例えば等しい出版年を持つ図書の集合を考えた場合，この集合の中では，貸出頻度が高い有用な図書と全く貸出されない図書とは区別されない。このことは，前章 D 節で述べた貸出頻度分布モデルの確率論的な解釈を考えれば，より明確になる。すなわち，貸出頻度分布モデルはその図書集合から「無作為に」一冊選んだときに，その図書の貸出回数が x 回である確率を与えるにすぎない。図書館利用を重視する観点から蔵書構築を行う場合には，貸出頻度の高い図書と低い図書とは区別して扱うことが望ましいと考えられるが，(19)～(22) 式ではそれは難しい。

この問題を解決するひとつの方法として，条件付き確率分布によるモデル $f(Z(t_2, h_2)|Z(t_1, h_1))$ を用いることが考えられる。例えば，初期時点から第 u 年までに m 回貸出された図書の第 u 年以降 v 年間の貸出頻度分布は，

$$f(Z(u, v)|Z(0, u)=m) \qquad (23)$$

で与えられる[16]。さらに $v\to\infty$ とすれば，(21)，(22) 式に対応したモデルを得ることができる。そして，このモデルを用い，例えば $m=0$ として計算すれば，未貸出図書の将来予測を行うことが可能になり，廃棄・別置に有用な情報が提供される。このように，(23) 式の m に適当な数値を設定すれば，貸出頻度の高い図書と低い図書とを区別して，「効用」を計算することができる。すなわち，貸出頻度分布モデルを用いれば，Brookes の「効用」よりも，さらに蔵書構築に便利な情報を得ることが可能である。

実際に，(10) 式に $t_2=u$, $h_2=v$, $t_1=0$, $h_1=u$ を代入して計算すると，

$$\pi(0, u, u, v)=\frac{1+(a/b)-e^{-au}}{1+(a/b)-e^{-a(u+v)}}$$

となる[16]。これから具体的に，(23) 式を計算することができる。さらに，このパラメターを用いた場合の (11) 式の平均は，$(k+m)(e^{-au}-e^{-a(u+v)})/[1+(a/b)-e^{-au}]$ で与えられる。

C. Morse の平均貸出回数予測モデルと貸出頻度分布モデル

Morse[1] は，貸出回数の経年的な変化を記述するモデルとして，

$$N(m)=\alpha+\beta m \qquad (24)$$

なる関係式を提案した。ここで，m は貸出回数，$N(m)$ は，貸出回数が m 回であった図書がその次の年に貸出される回数の平均であり，α, β はパラメターである。Kraft[21] や Beheshite and Tague[22] はさらに α を α_t に置き換えることによって，このモデルを，$N(m)=\alpha_t+\beta m$ と拡張した。実際のデータを用いて，これらのモデルから $N(m)$ を算出した結果，もし，$N(m)$ の値が十分に小さければ，貸出回数 m 回の図書を廃棄・別置の候補と考えることができる。

しかし，Burrell[15] は (24) 式が，貸出頻度分布モデル $f(Z(t_2, h_2)|Z(t_1, h_1))$ の特殊形に過ぎないことを示した。つまり (24) 式の $N(m)$ は，条件付き確率の期待値

$$N(m)=E[Z(t, 1)|Z(t-1, 1)=m]$$

として表現される。上の $N(m)$ の定義を見れば，この式が成立することは明らかである。

この期待値は貸出頻度分布モデル $f(Z(t, 1)|Z(t-1, 1)=m)$ を用いれば，容易に計算できる。したがって，Morse の平均貸出回数 $N(m)$ を予測するモデルも貸出頻度分布モデルによって表現され，さらに，貸出頻度分布モデルはその予測を平均としてだけではなく，確率分布 $f(Z(t, 1)|Z(t-1, 1)=m)$ として与えることがわかる。

実際，非定常ポアソン過程型貸出頻度分布モデル (10) 式に，$t_2=t$, $h_2=1$, $t_1=t-1$, $h_1=1$ を代入すれば，

$$\pi(t-1, t, 1, 1)=\frac{\phi e^{-a(t-1)}+1}{\phi e^{-a(t-1)}+\phi e^{-at}+1}$$

ただし，$\phi=(b/a)(1-e^{-a})$

となる[15]。これから，$f(Z(t, 1)|Z(t-1, 1)=m)$ の期待値を求めれば，

$$E[Z(t, 1)|Z(t-1, 1)=m]=(k+m)\frac{\phi e^{-at}}{\phi e^{-a(t-1)}+1}$$

である[15]が，$B_t=k\phi e^{-at}/(\phi e^{-a(t-1)}+1)$, $A_t=kB_t$ とおけば，

$$N(m)=A_t+B_t m$$

が得られる。これで (24) 式と同様のモデルが得られた。

D. 蔵書回転率と貸出頻度分布モデル

図書館の蔵書はいくつかの主題分野から構成されている。この各分野の蔵書の量的な比率が，その図書館にとって適切であるかどうかは，蔵書構築にとって重要な問題である。例えば，活発に利用される分野の蔵書の全体

に対する比率が小さければ，その分野の蔵書数を増やして強化することが必要になろう。そこで，利用データからこの蔵書の比率を評価する指標として，「蔵書回転率」が提案されている。

蔵書回転率は，Ei/Ci で定義される[23]。ここで，Ei は分野 i の貸出回数，Ci は分野 i の蔵書数である。すなわち，蔵書回転率とは蔵書一冊あたりの貸出回数を意味している。したがって，貸出頻度分布モデル $f(Y_t)$ を用いれば，蔵書回転率は，

$$\sum_{Y_t=0}^{L} Y_t f(Y_t), \quad (t \text{ は固定})$$

と表現される。つまり，蔵書回転率とは $f(Y_t)$ の期待値に相当する（ただし，これは1年間の貸出を問題とした場合であり，複数年ならば $f(X_t)$ あるいは $f(Z(t, h))$ を用いることになる)。

したがって，貸出頻度分布は蔵書回転率よりも，蔵書の貸出の状況についての情報をより多く含んでいることになる。例えば，蔵書回転率が小さい場合，それが蔵書の欠陥によるものか，それともその分野の貸出が不活発であることによるものかのどちらであるかに関しては蔵書回転率からだけでは判断できない。一方，貸出頻度分布ならば，この種の判断を分布全体の形状等を見て行うことが可能である。例えば，貸出頻度分布が右裾（確率変数 x が大きい部分）を長く引くにもかかわらず，平均が小さければ，その蔵書は欠陥のある可能性があり，複本購入あるいは廃棄・別置を考えることが示唆される。なぜなら，この形状は少数の図書に利用が集中する一方で，大部分の図書が利用されないという状況を示しており，この状況はアベイラビリティや書架スペースの効率等の点から望ましくないと考えられるからである。一方貸出頻度分布が右裾を長く引かず，平均も小さい場合，その分野の貸出そのものが不活発であると判断できる。

また，蔵書回転率とは別に，Bonn[24]が提案した“use factor（利用係数)”という指標がある。これは，$(Ei/E)/(Ci/C)$ で定義される。ここで，E は全貸出回数，C は全蔵書数である。したがって，この指標は，分野 i の貸出回数の全貸出回数に対する比率を分野 i の蔵書数の全蔵書に対する比率で割った値を意味している。ところが，河井[25]が示しているように，この指標と蔵書回転率との間には，

$$\frac{Ei/E}{Ci/C} = \frac{Ei/Ci}{E/C} \qquad (25)$$

という関係がある。右辺の分母は，蔵書全体での蔵書回転率であり，これもやはり貸出頻度分布モデルで表現することができる。したがって，“use factor”も貸出頻度分布モデルによって表されることがわかる（なお，(25)式右辺を河井[25]は「回転率指数」と呼んでいる)。

E. ブラッドフォードの法則と貸出頻度分布モデル

ある主題に関して，各雑誌が掲載している関連文献の数でランキングを行った場合，その順位 n と順位 n までの雑誌の文献数の累積総数との間にブラッドフォードの法則が成り立つことが知られている。このブラッドフォードの法則に対しては，様々な数式表現が提案され，その妥当性の検証が試みられている。

ブラッドフォードの法則は，少数の雑誌に関連文献が集中する一方，関連文献を非常に僅かしか掲載していない雑誌が多数存在するような，非常に偏りのある分布形態を表現している。この偏りのある分布形態は蔵書構築にとって重要である。なぜなら，各雑誌の価値が均等な状態に比べて，このような偏った状態では，価値が大きな少数の雑誌を優先的に図書館が所有することにより，より少ない経費・スペースでより多くの利用者の要求を充足することができるからである。

しかし，ブラッドフォードの法則が掲載文献数による順位づけだけでなく，被引用回数や利用回数によって順位づけを行った場合でも成立することが知られている。また，雑誌ではなく，図書の貸出の場合にもあてはまることも報告されている[26]。

貸出頻度分布が度数分布であるのに対し，ブラッドフォードの法則は順位分布の一種である。順位分布は基本的に，各対象（雑誌，図書など）をその大きさ x（掲載文献数，貸出回数など）の順で並べたとき，その順位 n を独立変数として，$x=r(n)$ と表現される。特に，ブラッドフォードの法則の場合は，順位 n までの x の累積総数を問題とする。すなわち，$\sum r(n)$ である。順位分布を用いると，利用回数の多い雑誌上位100誌だけを受け入れた場合，何%の要求を充足できるか，などの計算を行うことができ，蔵書構築にとって便利な場合がある。

そこで，貸出頻度分布モデルをこの順位分布の形式に数学的に変換することが考えられる。正規化された度数分布（貸出頻度分布モデルに相当）を $f(x)$ と記す。順位分布から正規化された度数分布への変換は，

$$f(x) = -\frac{1}{N}\frac{d}{dx}r^{-1}(x)$$

であり[27]，比較的容易である（N は対象の総数）。しかし，ここで問題となるのは，度数分布から順位分布へという，逆方向の変換である。正規化されない度数分布を $g(x)$ と記す。つまり $f(x)=g(x)/N$ である。さらに，最大の x を L と記す（例えば，貸出回数の最大値）。ある x を持つ対象の順位は，$\int g(x)dx=G(x)$ を用いると，

$$n=G(L)-G(x)+1$$

と記すことができる。したがって，$G(x)=G(L)+1-n$。ここで x について解ければ，

$$x=G^{-1}[G(L)+1-n]=r(n) \qquad (26)$$

となる。ただし，これはブラッドフォードの法則の形式ではなく，通常の順位分布の形式であるから，ブラッドフォードの法則の形式を得るにはさらに (26) 式を積分する必要がある。

一方，厳密なブラッドフォードの法則の形式ではないが，x 回以上の貸出回数の図書が全体の貸出回数の何％を占めるかを求める関数は，(26) 式よりも容易に書き表すことができる。この関数を $\eta(x)$ と記す。$\eta(x)$ は貸出回数の場合であるから，貸出頻度分布モデル $f(x)$ を用いると，

$$\begin{aligned}\eta(x)&=\sum_{i=x}^{L} if(i)/\sum_{i=0}^{L} if(i)\\ &=1-\sum_{i=0}^{x-1} if(i)/\sum_{i=0}^{L} if(i) \qquad (27)\end{aligned}$$

となる[28]。

この (26)，(27) 式より，貸出頻度分布モデルによってブラッドフォードの法則のひとつの表現を得ることができた。すなわち，ブラッドフォードの法則は貸出頻度分布モデルによって表現可能であることが示せた。さらに $f(x)$ の代わりに $f(x,t)$ を用いれば，経年変化を記述するブラッドフォードの法則の数式表現も得ることもでき，ブラッドフォードの法則の拡張が可能になる。

実際，Burrell[28] によって，負の二項分布の特殊形 ($k=1$) である幾何分布を $f(x)$ とする場合の $\eta(x)$ が考察されている。ただし，Burrell は，x を貸出回数とするのではなく，貸出回数の大きい順に図書を並べた場合の上位 x% の図書に対する貸出回数の割合 $\eta(x)$ を求めた。これは (27) 式よりもブラッドフォードの法則にさらに近い表現である。これを $\eta(x')$ と記す。そして，幾何分布を $f(x)$ とする場合のこの関数が実際に，

$$\eta(x')=x'\left(\frac{p}{\log(1-p)}\log x'+1\right)$$

として求められている[28]。しかし，負の二項分布自体についての $\eta(x')$ はまだ具体的に示されておらず，今後の課題である。

IV. 貸出頻度分布モデルの応用可能性

第Ⅲ章では，蔵書構築のためのいくつかの尺度・手法・モデルが，貸出頻度分布モデルによって表現されたり，あるいは取って代わられることを示した。これによって，貸出頻度分布モデルが蔵書構築のための計量的手法のひとつの基礎であることが明らかになったと言えよう。

次に本章では，貸出頻度分布モデルが実際にいくつかの計量的手法に応用され，有益な結果をもたらすことについて論じる。特にここでは，ひとつの事例として，Morse[1] によって開発された複本購入問題等のための待ち行列モデル，および Trueswell[29] による廃棄・別置の手法に貸出頻度分布モデルを応用することの有用性を見る。

A. Morse の待ち行列モデルへの貸出頻度分布モデルの応用可能性

Morse[1] は，複本購入問題や貸出期間の設定問題にオペレーションズ・リサーチの分野で発達した待ち行列理論を応用することを試みた。待ち行列理論は，基本的には，窓口でのサービスに関して，客の到着やサービス時間に応じて窓口の数等を数学的に決定する理論である。Morse[1] は窓口を図書，サービスを貸出，サービス時間を貸出期間と置き換えることにより，待ち行列理論を複本購入問題に応用することに成功した。ここで，客が多すぎるためにひとつの窓口で客に対応しきれない場合に窓口を増やすことが，複本の購入に対応する。

この Morse のモデルのひとつの欠点は，一冊の図書のある一時点の状況しか記述できない点にある。

例えば，Morse のモデルで最も単純な関係式として，

$$R=\lambda\omega/(\lambda+\omega)$$

がある。ここで，R はある一冊の図書が 1 年間に貸出される回数，λ は 1 年間にその図書を利用したいと望んだ人の延べ人数，ω は平均貸出期間（単位は 1 年）の逆数である。この式は，書架上に図書がなかった場合，その利用者は貸出予約を行わずに，そのまま立ち去るこ

とを前提としている。この式は，

$$\lambda=R\,\omega/(\omega-R)=[\frac{1}{R}-\frac{1}{\omega}]^{-1} \qquad (28)$$

と変形できるので，その図書の貸出回数 R と平均貸出期間の逆数 ω をデータより算出できれば λ が求められる。そして，その結果，その図書を利用したいと望みながらも，その図書が貸出中のために利用できなかった人の数を推定することができ，複本購入などによるアベイラビリティ改善への手がかりとなる。

しかし，これはあくまで一冊の図書についての結果にすぎないし，しかも obsolescence が考慮されていないため，この λ の推定値が次年度も正しいかどうかはわからない。したがって，(28) 式だけでは，実際の蔵書構築には不十分である。アベイラビリティが蔵書全体ではどうなっているのか，あるいは，利用できないための不満が将来どの程度生じるかなどが予測できなければ，実際の複本購入には十分に役立たない。

この予測を可能にするためには，Morse のモデルに貸出頻度分布モデルを組み込むことが考えられる。まず，(28) 式を $\phi(R)$ とおく（ω は定数とする)。R は貸出回数であるから，データから蔵書全体の貸出頻度分布モデル $f(R, t)$ のパラメターを推定することができる。そこで，第 t 年における λ の蔵書全体での平均は，R を変数 x に置き換えて，

$$E(\phi(x))=\Sigma\phi(x)f(x, t) \qquad (29)$$

によって計算することができる。さらにこれに蔵書数 C をかければ，1 年間にその図書を利用したいと望んだ人の蔵書全体での延べ人数を算出できる。また，利用を望んだが，貸出中のために利用できなかった人の蔵書全体での総数 W は，

$$W=C(\Sigma\phi(x)f(x, t)-\Sigma\, xf(x, t))$$

として求めることができる。この W はあくまで延べ人数であり，他人の貸出中のため書架上に図書がなかった場合の総回数に相当する。(28) 式は，本来，ある一冊の図書のある一時期の利用を対象とするものであったが，(29) 式のように貸出頻度分布モデルと組み合わせることにより，蔵書全体の問題，あるいは，将来的な予測の問題にまで拡張することが可能になった。

例として，慶應大学日吉情報センターにおける，1982 年度に出版された図書に関する 1983 年度の貸出頻度分布のデータ（第 1 表参照）で計算を行ってみる。ただし，第 1 表の貸出についての平均貸出回数のデータがないので，これを 2 週間と仮定する。1 年間を 52 週と考えることにすると，2 週間は 2/52 年。したがって $\omega=26$。この数値を用い，(28) 式を各 x について順次計算すると，$E(\phi(x))$ は約 1.24 になる。$E(x)$ は 0.96 であり，蔵書数は 9480 であるから，$W=2654.4\cdots$となる。したがって，総貸出回数は 9434 回であるから，貸出が成功する確率は約 78%(=9434/(9434+2654)) と計算できる。この数値から 5 回のうち約 1 回は貸出中のため書架上に図書がないことがわかる（ただし，これはあくまでも (29) 式の実際の用い方を例示するものにすぎず，貸出予約が行われないという前提と，平均貸出回数

第 1 表　慶應大学日吉情報センターにおける 1982 年に出版された図書についての 1983 年度の貸出頻度分布

貸出回数	冊　数
0回	5542
1	1876
2	826
3	475
4	283
5	170
6	92
7	72
8	44
9	24
10	36
11	16
12	9
13	5
14	2
15	2
16	3
17	0
18	1
19	0
20	0
21	1
合　計	9480
総貸出回数	9434
平　均	0.96

出典：岸田ほか[12]，p. 34

が 2 週間という仮定に基づく計算である。以下の計算も同様)。

次に，仮に貸出回数 10 回以上の図書に対してそれぞれ複本を一冊用意していたとして，W の値がどう変わるかを試算してみよう。各 $\lambda=\phi(x)$ は前の計算と変わらないとする。元の図書と複本とを合わせて一冊(一単位)と考え，貸出回数 x 回であった図書に複本を一冊購入した場合にその図書と複本の両方に対して行われる貸出回数を $\zeta(x)$ で表す。この $\zeta(x)$ は Morse のモデルより，

$$\zeta(x)=\frac{\phi(x)(\omega^2+\phi(x)\omega)}{\omega^2+\phi(x)\omega+\phi(x)^2/2}$$

と計算できる。したがって，貸出回数 10 回以上の複本を持つ図書に対する貸出回数の総計 Tc は，

$$Tc=\sum_{x=10}^{21} C\zeta(x)f(x,t)\,, \quad (t \text{ は固定})$$

となる。実際に上の例で計算してみると，Tc は約 1352 となる。貸出回数 0 回から 9 回までの図書の貸出回数の総数は，第 1 表より，8559 回であるから，複本を購入した場合の総貸出回数は，1352+8559=9911 となる。したがって，$W=1.24\times 9480-9911=1844.2$ となり，先の W の値の約 69% に減ることがわかる。すなわち，わずか 76 冊の複本 (全蔵書の 0.8%) によって元の W の約 30% の改善が達成されることが示される。

この例が示すように，待ち行列モデルと貸出頻度分布モデルとを組み合わせることによって，複本購入に対して非常に有用な情報を得ることができる。ただし，上の例には貸出予約が行われないという条件がついていた。しかしながら，Morse[1] には貸出予約する場合，あるいは予約の人数に制限のある場合などさまざまなケースが想定されて，モデルが作成されている。したがって，これらのモデルに貸出頻度分布モデルを組み込めば，応用範囲は広がる。

さらに，貸出頻度分布モデル $f(Z(t_2,h_2)|Z(t_1,h_1))$ を用いれば，複本の貸出回数の将来予測を行うこともできる。例えば，第 t 年次に m 回以上の貸出が行われた図書を複本購入の候補とするならば，

$$f(Z(t,\infty)|Z(t-1,1)\geqq m)$$

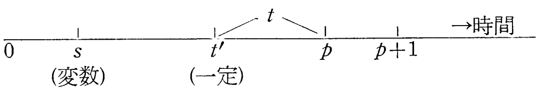
によって，その将来の貸出回数を予測することが可能である。そして，$\phi(x)$ や $\zeta(x)$ を用いて，複本の効果をある程度予測することができる。特に貸出頻度分布モデルとして，非定常ポアソン過程型モデルを用いれば，このモデルには obsolescence の要因が含まれているため，より現実に即した予測が可能になる。第 I 章でも述べたように，従来の Morse のモデルの欠点は，時系列的な予測を行えないことであったが，貸出頻度分布モデルの予測の能力をこれに加えることによって，この欠点が克服される。

以上のように，貸出頻度分布モデルを Morse の待ち行列モデルに適用することにより，複本購入問題に対してより有用な手法を手に入れることができる。

B. Trueswell の最後の貸出日と貸出頻度分布モデル

貸出データに基づく図書の廃棄・別置の方法として，Trueswell[29] の「最後の貸出日 (last circulation date)」や Slote[30] の「在架期間 (shelf time period)」に基づく手法がある。これらはいずれも，現時点と，最も最近その図書が貸出された時点との間の時間間隔を廃棄・別置の目安としている (ただし「最後の貸出日」はその調査時点に一番近い過去の貸出日付を指し，「在架期間」はその図書が書架上に存在した時間を意味するから，両者は厳密には一致しない)。実際には，これらの方法では，ある時点で貸出データを調べて，その貸出データに最近 t 年間に一度も貸出されなかった図書が含まれていない場合に，「最近 t 年間以上貸出されなかった図書を廃棄する」のような決定を行う (未貸出の図書はその出版年あるいは受け入れ年で「最後の貸出日」を代用する)。

この廃棄・別置の候補となる図書数を貸出頻度分布モデルで表現してみる。前章 A 節同様，第 s 年に出版された (受け入れられた) 図書に関する貸出頻度分布モデルを $f_s(x,t)$ と書き，第 s 年に出版された (受け入れられた) 図書に関する蔵書数を C_s とする。そして，時点 p から $p+1$ ($p+1$ 年次) における貸出データから「最近 t 年間以上貸出されなかった図書を廃棄する」と決めたとする。すると時間軸は，



となる。ここで，$t=p-t'$。この s, t', p は，各 $f_s(x,t)$ で共通の定数であることに注意せよ。したがって，例えば，$t'=10$, $p=20$ の場合，出版年 $s=5$ の図書の時点 p における貸出頻度分布は $f_{10}(x,20-5)=f_{10}(x,15)$ によって表現される。これらの記号を用いると，廃棄される図

書の数は，

$$\sum_{s<t'} C_s f_s(Z(p-s,1)=0|Z(t'-s,p-t')=0) \quad (30)$$

によって求められる。式中の条件付き確率は，過去 t 年間（$p-t'$ 年間）に一度も貸出されなかった図書が調査時点 p から $p+1$ の間（$p+1$ 年次）にも貸出されなかったことを意味する。また，式中の和を $s<t'$ についてとる理由は，先に述べたように，最近 t 年間未満に出版あるいは受け入れられた図書が未貸出の場合は，その出版年あるいは受け入れ年で「最後の貸出日」を代用するために，$s>t'$ である図書は除いて考える必要があるためである。

Trueswell あるいは Slote の手法は，廃棄・別置する図書を決定するだけに留まる。しかし，(30) 式を利用して貸出頻度分布モデルを応用すれば，廃棄・別置した図書の将来の貸出回数を確率的に予測することができる。すなわち，これらの図書が将来貸出される確率は，各出版年別に，

$$f_s(Z(p+1-s,\infty)|Z(t'-s,p+1-t')=0)$$

によって求めることができる。この確率分布の $x=0$ の部分が大きければ，廃棄・別置の決定は正しい可能性が大きいと言える。

このように，貸出頻度分布モデルを応用することによって，より精緻な廃棄・別置を行うことができる。

V. おわりに

本論文では，蔵書構築に関して現在までに提案されている手法・尺度・モデルのいくつかが貸出頻度分布モデルから導かれ，さらに貸出頻度分布モデルがそれらと同等以上の機能を持っていること，また，貸出頻度分布モデルが，いくつかの手法に応用可能であることが示された。これにより，貸出頻度分布モデルが，蔵書構築のための計量的な手法の基礎となるモデルであることが明らかになった。確かに，貸出頻度分布モデルがすべての尺度・手法・モデルを表現できるわけではない。しかし本論文で取り上げた蔵書構築のための計量的手法は，現在有用であるとされている図書の貸出データに基づく手法のかなりの部分を網羅していると考えられる。そこで，貸出頻度分布モデルが基礎的なモデルのひとつであることは十分示されたと言えよう。

今後の課題としては，具体的な貸出頻度分布モデルの精緻化が挙げられる。本論文では，貸出頻度分布モデルの具体例として，非定常ポアソン過程型モデルを取り上げてきた。そして，このモデルが現実に対してある程度妥当性を持っているために，貸出頻度分布モデル $f(x,t)$ が実際に設定できるものとして議論を進めてきたのであった（第II章参照）。しかし，負の二項分布が貸出頻度分布を記述することはかなりのデータにより実証されているものの，非定常ポアソン過程型モデルが経年変化を記述するかどうかについては，まだデータによる実証が十分でない。しかも，非定常ポアソン過程型モデルの各年次分布 $f(Y_t)$ をより一般化したモデルも提案されている[12]。ただし，この一般化されたモデルは $f(Y_t)$ についてのみであり，$f(X_t)$, $f(Z(t,h))$ に関してはまだ一般化されたモデルが示されていないので，今回は貸出頻度分布モデルの具体例としては用いなかった。このような貸出頻度分布モデル $f(x,t)$ の精緻化は，さらに進める必要がある。

この貸出頻度分布モデルの精緻化が実現した場合，本論文で示したいくつかの関係式を用いることにより，他の手法やモデルも同時に精緻化されることになる。

1) Morse, Philip M. Library effectiveness: a system approach. Cambridge, MIT Press, 1968. 207 p.
2) この欠点に対する修正案は，一応，
Chen, Ching-Chin. Applications of operations research models to libraries: a case study of the use of monographs in the Francis A. Countaway library of medicine, Harvard University. Cambridge, M.I.T. Press, 1976, 212 p. で提示されている。
3) 第III章E節で述べるように，ブラッドフォードの法則は本来は，雑誌の掲載文献数についての数量的な関係を記述するものであるが，被引用回数や利用回数にも応用されている。
4) Burrell, Quentin L. A simple stochastic model for library loans. Journal of Documentation. Vol. 36, No. 2, p. 115-132 (1980)
5) Burrell, Quentin L.; Cane, Violet R. The analysis of library data. Journal of the Royal Statistical Society, Series A. Vol. 145, p. 439-463 (1982)
6) Wall, T. Letter to the editor. Journal of Documentation. Vol. 36, No. 4, p. 343-344 (1980)
あるいは，
Chatfield, C. Discussion of the paper by Mr. Burrell and Professor Cane. Journal of the Royal Statistical Society, Series A. Vol. 145, p.

463-471 (1982)
7) Bagust, A. A circulation model for busy public libraries. Journal of Documentation. Vol. 39, No. 1, p. 24-37 (1983)
8) Burrell, Quentin L. A note on ageing in a li. brary circulation model. Journal of Documentation. Vol. 41, No. 2, p. 100-115 (1985)
9) Gelman, E.; Sichel, H. S. Library book circulation and the beta-binomial distribution. Journal of the American Society for Information Science. Vol. 38, No. 1, p. 4-12 (1987)
10) Sichel, H. S. A bibliometric distribution which really works. Journal of the American Society for Information Science. Vol. 36, No. 5, p. 4-12 (1985)
11) Cocks, T. M.; Brookes, B. C. Sichel's unification of bibliometric frequency distributions. Journal of Information Science. Vol. 12, No. 1, p. 45-51 (1986)
12) 岸田和明ほか．大学図書館における図書の貸出頻度についての確率過程モデルの検討：負の二項分布を中心として．Library and Information Science. No. 25, p. 25-39 (1987)
13) 岸田和明．ビブリオメトリックスの現象を記述する確率分布の比較．情報の科学と技術．Vol. 40, No. 6, p. 427-437 (1990)
14) ただし，岸田[13]は $x=0$ の部分を切断した片側負の二項分布について調べた。また，その片側負の二項分布は特に文献の利用に関するデータに対して適合した。
15) Burrell, Quentin L. A second note on ageing in a library circulation model: the correlation structure. Journal of Documentation. Vol. 42, No. 2, p. 114-128 (1986)
16) Burrell, Quentin L. A third note on ageing in a library circulation model: applications to future use and relegation. Journal of Documentation. Vol. 43, No. 1, p. 24-45 (1987)
17) Gosnell, Charles F. Obsolescence of books in college libraries. College and Research Libraries. Vol. 5, No. 2, p. 115-125 (1944)
18) Burton, R. E.; Kebler, R. W. The "half-life" of some scientific and technical literatures. American Documentation. Vol. 11, No. 1, p. 18-22 (1960)
19) Line, M. B.; Sandison, A. 'Obsolescence' and changes in the use of literature with time. Journal of Documentation. Vol. 30, No. 3, p. 283-350 (1974)
20) Brookes, B. C. Obsolescence of special library periodicals: sampling errors and utility contours. Journal of the American Society for Information Science. Vol. 21, No. 5, p. 320-329 (1971)
21) Kraft, D. H. A comment on the Morse-Elston model of probabilistic obsolescence. Operations Research. Vol. 18, No. 6, p. 1228-1233 (1970)
22) Behesite, Jamshid; Tague, Jean M. Morse's Markov model of book use revised. Journal of the American Society for Information Science. Vol. 35, No. 5, p. 259-267 (1984)
23) 森耕一．蔵書構成の適否をはかる一方法．図書館界．Vol. 23, No. 4, p. 161-163 (1971)
24) Bonn, George S. Library self-surveys. Library and Information Science. No. 9, p. 115-121 (1971)
Bonn, George S. Evaluation of the collection. Library Trends. Vol. 22, No. 3, p. 265-304 (1974)
25) 河井弘志．アメリカにおける図書選択論の学説史的研究．東京，日本図書館協会，1987. p. 153.
26) Bulick, Stephen. Book use as a Bradford-Zipf phenomenon. College and Research Libraries. Vol. 39, No. 3, p. 215-219 (1978)
27) 岸田和明．ビブリオメトリクスの法則間の類似関係から導かれる Bradford の数式表現．Library and Information Science. No. 26, p. 55-65 (1988)
28) Burrell, Quentin L. The 80/20 rule: library lore or statistical law? Journal of Documentation. Vol. 41, No. 1, p. 24-39 (1985)
29) Trueswell, Richard W. A quantitative measure of user circulation requirements and its possible effect on stack thinning and multiple copy determination. American Documentation. Vol. 16, No. 1, p. 20-25 (1965)
30) Slote, Stanley J. Weeding library collections—II. Littleton, Colorado, Library Unlimited, 1982. 198 p.