

## 拡散近似——その考え方と有用性——

木村 俊一

### 1. まえがき

待ち行列現象は、直接あるいは間接に、さまざまなレベルで私たちの日常生活とかかわりあっている。誰にとっても待つことは決して愉快なことではないし、待たないに越したことはない。待ち行列理論は、本来こういった素朴な願いに応えるべく誕生してきたし、またそうでなければならぬはずである。この意味では、現在の待ち行列理論は、あまりに数学的になりすぎたように感じられる。この数学化の傾向は否定すべきことでもないし、ある意味では歓迎すべきことであるが、その一方で、数学モデルのための待ち行列理論という印象を与えてはいないだろうか。解析のために都合の良い仮定を並べて解いたところで、どれほどの意味があるだろう。また、たとえ解けたとしても、その解が母関数、Laplace変換等の複雑な形や、大規模な数値計算を必要とするものであったとしたら、多くの人は失望の色を隠せないだろう。こう考えてみると、ORのたいていの教科書の中で、待ち行列の章が最後のほうに位置しているのも、うなずける気がしてくる。

待ち行列における近似モデルの重要性は、こういった状況の中で、ますます高いものになっている[23]。使われる数学の程度が高いか低いかは問

題ではない。要は、得られる結果が使いやすいか否かである。このためには、モデルの厳密さを犠牲にすることも止むを得ないだろうし、またそうしなければ、残された多くの待ち行列問題が解けないことは、ほとんどの研究者が承知している“事実”である。

拡散近似(diffusion approximation)は、1960年代後半から展開をとげてきた待ち行列の近似モデルの1つである。簡単に言ってしまうと、拡散近似とは待ち行列特性量、たとえば系内客数、仮りの待ち時間等の確率過程を、適当な拡散過程で近似することに他ならない。数学的には、出生死滅過程(birth and death process)の状態空間を連続化することに相当している。こういった近似は、何も待ち行列だけにすぎず、数理生物学等の分野で一種の連続近似として、ずっと以前から用いられていたものである[9]。このように拡散近似は、考えようによっては至極“単純”とも取れる近似ではあるが、その正当性を評価するにあたって、意外に微妙で複雑な側面が現われてくる。以下では、複数窓口待ち行列の系内客数の過程を例としてとりあげ、『なぜ拡散近似が妥当なのか?』について、確率測度の弱収束の概念を用いて厳密に示すことにしよう。これにより、拡散近似のもつ問題点や近似の有用性がおのずと明らかになるはずである。

## 2. 近似の正当性

拡散近似が真に正当な近似なのかどうかを示すためには、近似の対象である待ち行列特性量の確率過程と対応する拡散過程との間に“近さ”の概念を導入する必要がある。普通こういった場合、それぞれの過程の分布の間の“近さ”が目安になり、中心極限定理が重要な役割をはたしているが、ここではその役割を確率測度の弱収束極限定理にはたしてもらおう。弱収束の概念を用いることで、確率過程間の“近さ”を直接的に測ることが可能になり、より精密な評価をくだすことができる。

まず、弱収束の概念を簡単に述べることにする(詳細は、[1] 参照)。近似の対象としている確率過程の見本関数は、必ずなんらかの不連続性をもっている。たとえば、複数窓口待ち行列の系内容数過程の見本関数は、図1に示すように、客の到着とサービス終了前後で不連続に変化している。この不連続性は、仮りの待ち時間等の過程にも同様に見出すことができる。このような不連続点をもつ確率過程を扱うために、関数空間  $D[0, 1]$  ( $\equiv D$ ) を導入しよう。 $D[0, 1]$  は、左極限值をもつ閉区間  $[0, 1]$  上のすべての右連続関数の作る空間で、対象とする確率過程は、適当な時間軸の変換によって、以下に定義する  $D$  における確率関数 (random function) の系列として生成されるものと解釈できる。さらに、 $D$  の位相的ボレル集合体、すなわち、開集合を含む最小のボレル集合体を  $\mathcal{D}$  で表わすことにする。このとき、 $\mathcal{D}$  上の確率測度の弱収束は次のように定義される。

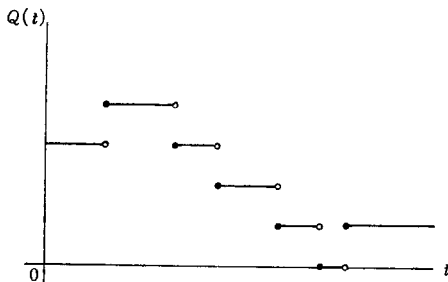


図1 系内容数過程

## 特集に当って

茨城大学 森 雅夫

最近の待ち行列は変わってきていると思う。ひところのように、モデルを解析的に解いて、ラプラス変換を求めてよしとやり方は、はやらない。なんとか答を数値化して見せる。行列長の分布や待ち時間分布まで出してしまおう。これもひとえにコンピュータ進歩の影響だろう。これは、待ち行列にかぎらず確率モデル全般の傾向でもあるようで、Computational Probability や Algorithmic Method in Probability などの題目の論文集も始めているほどである。

高橋氏には、強くなった数値解法ということで、その辺の状況をまとめていただく予定であったが、氏の手の届く範囲で調べられた結果、数値計算例は多くなってきたが、解を求めるためのアルゴリズム的手法を意識したものは未だ少ないという。そこで氏には、マルコフ型モデルに対するアルゴリズム的手法を解説していただく。

トラフィック密度  $\rho$  が1以上のとき、待ち行列過程そのものが拡散過程で近似できることは以前から知られている。木村氏の論文は、 $\rho < 1$  のときでも工夫すれば、この拡散モデルによる近似はかなり使えることを教えている。

橋田氏には、ここ数年、大きな話題であったネットワーク型待ち行列について、モデルのいろいろとその解法を紹介していただく。逆瀬川氏には、やはり1つの数値解法であるシミュレーションについて、データのとり方、まとめ方などに関して最近の理論的成果をまとめていただいた。

通信、コンピュータなどへの応用例を紹介するという手もあったが、それらはいずれ別のタイトルの下で特集が組まれるだろうと期待して、ここではモデルを精しくあるいは近似的に解くことに関しての理論の発展状況を紹介するにとどめた。

**定義 1**  $\mathcal{D}$  上の確率測度  $P_n$  と  $P$  が、 $D$  上の任意の有界で連続な実数値関数  $f$  に対して、

$$\lim_{n \rightarrow \infty} \int_D f dP_n = \int_D f dP \quad (1)$$

を満たしているならば、 $n \rightarrow \infty$  のとき、 $P_n$  は  $P$  に弱収束する ( $P_n$  converges weakly to  $P$ ) といいい、 $P_n \Rightarrow P$  と書く。

この定義は、確率測度間の収束について述べているだけで、確率過程間の収束については何も述べていない。そこで、次の定義が必要になる。

**定義 2** 確率関数  $X$  とは、確率空間  $(\Omega, \mathcal{F}, P)$

† Kolmogorov [20] の定義した確率関数は、一価関数  $X$  に対する分布  $PX^{-1}$  をさし、この定義とは異なる。

から  $D$  への可測写像のことであり,  $(D, \mathcal{D})$  上に分布  $P=PX^{-1}$  をもっている.

上の2つの定義より, 確率過程の間の弱収束の概念は本質的には次の定義によって与えられる.

**定義 3**  $\{X_n\}$  を確率関数の系列であるとする.  $X_n$  の分布  $P_n$  が, ある確率関数  $X$  の分布  $P$  に弱収束するならば,  $X_n$  は  $X$  に弱収束するといひ,  $X_n \Rightarrow X$  と書く.

以上述べた弱収束の概念とそれにもとづく弱収束極限定理の結果が, 拡散近似にいかん反映しているかを示す一例として,  $GI/G/s$  待ち行列の系内客数過程をとりあげることにしよう. 最初に, このシステムを特徴づけるパラメータと記号を明確にしておこう. システムへの客の到着時間間隔は, 独立で同じ分布にしたがう確率変数で, 平均  $1/\lambda$  と有限の分散  $\sigma_a^2$  をもっているものとする.  $s$  個の窓口でのサービス時間は, いずれも独立で同じ分布にしたがひ, おのおのが平均  $1/\mu$  と有限の分散  $\sigma_s^2$  をもっているものとする. システムのトラフィック密度(traffic intensity)を  $\rho=\lambda/s\mu$  で表わし, さらに, 時刻  $t$  でのサービス中の客を含めた系内客数を  $Q(t)$  で表わすことにする. このとき, 系内客数過程  $\{Q(t); t \geq 0\}$  に対応して,  $D$  の確率関数  $Q_n$  を次式で定義する.

$$Q_n = \frac{Q(nt) - (\lambda - s\mu)nt}{a\sqrt{n}}, \quad 0 \leq t \leq 1 \quad (2)$$

ただし,  $a = \lambda^3 \sigma_a^2 + s\mu^3 \sigma_s^2$ . Iglehart と Whitt [13] は, 不安定な (unstable), すなわち,  $\rho \geq 1$  であるシステムに対して, 次の2つの重負荷極限定理 (heavy traffic limit theorems) を示した.

**定理 1**  $\rho > 1$  ならば,  $Q_n \Rightarrow \xi$  が成り立つ. ただし,  $\xi$  は標準ウィーナー過程を表わす.

**定理 2**  $\rho = 1$  ならば,  $Q_n \Rightarrow f(\xi)$  が成り立つ. ただし,  $f$  は  $D$  からそれ自身への写像で,  

$$f(x)(t) = x(t) - \inf\{x(u), 0 \leq u \leq t\}, \quad 0 \leq t \leq 1 \quad (3)$$

で定義される.

(3)式で定義される写像  $f$  は, 容易に確かめら

れるように, 原点  $x(\cdot) = 0$  での通過できない境界の役割を果たしているので,  $f(\xi)$  は原点に反射壁境界をもつウィーナー過程を表わし,  $|\xi|$  と同じ分布をもっている. また, 定理1は, 定義3から本質的には中心極限定理の内容とまったく等価である. これら2つの極限定理から, 不安定な  $GI/G/s$  待ち行列の系内客数過程は, 適当な拡散パラメータをもつ自由 (free) または反射 (reflected) ブラウン運動過程で『近似』できることがわかる. ここで『近似』としたのは, これらの極限定理が十分長い時間経過した後の系内客数過程の漸近的な特性を示したものにすぎないからで, 不安定なシステムであっても状態空間がまったく異なる以上, しょせん, 近似でしかあり得ないのである. まして安定な (stable), すなわち,  $\rho < 1$  であるシステムに対しては, とてこのような極限定理の成立を望むことはできない. もしなんらかでも拡散近似の理論的根拠を見出そうとするならば, それは重負荷 (heavy traffic) 時を除いて他には考えられない. つまり,  $\rho \geq 1$  の場合に定理2の結果を拡大解釈して, 安定なシステムの系内客数過程も, 反射ブラウン運動過程で近似できるものとするのである. この大雑把ともいえる解釈が, 拡散近似の骨組みを支えているといっても過言ではないだろう. しかし, このことは近似の正当性を決して否定するものではなく, むしろ近似の結果によってその良否が判定されるべき性質のことのように思える.

### 3. 定式化における問題点

待ち行列の特性量を解析するに当って, 定常状態におけるそれが対象となるのが普通である. これは, 定常状態がそれ自体特に重要であるというよりは, 過渡状態の解析が極端にむずかしいという理由による部分が多い. 拡散近似に関しても, 極限定理は特性量の漸近的な特性しか保証してくれないので, 過渡状態へ適用すること[22]には, 単純には受け入れられない点がある. したが



似公式を得ることができるからである。このように、拡散パラメータの決め方如何によっては、解の扱いやすさ、さらには解の精度にまで大きな影響が現われてくるので、これらを比較したうえでいずれを取るかを定める必要がある。一般に、状態に依存しない拡散パラメータは、Heyman が与えた再生定理を用いる方法 [12] でほぼ一意に決定できるが、状態に依存する場合には、必ずしも一意に決まらない。しかし、このことは解の精度を高めるための自由度が残されているという意味で、むしろ好ましいことといえる。対象となるシステムの特長性をうまく利用できるかどうか、決め方のポイントになるだろう。

### 3.2 境界条件

待ち行列特性量の取り得る値には、たいていなんらかの制約がある。系内客数についていえば、非負の値しか取り得ないし、場合によっては上限値が定まっているときもある。他の過程についても同様である。したがって、これらを近似する拡散過程についても当然同じ制約が課せられることになる。その代表的なものが定理 2 で示された原点における反射壁境界だが、この境界は、重負荷時を除いてはあまり適当でないことが数値的に確かめられている [17, 第 3 章]。これは、原点における確率質量 (probability mass) の大きさに起因しているため、軽負荷時にはこの影響が無視できなくなってくる。つまり、システムが空である確率を考慮せざるを得なくなるわけである。この軽負荷時の反射壁境界の欠陥を改善する一般的な方法は知られていないが、システムの特長性を利用するいくつかの方法がある。その 1 つは、客がポアソン到着する場合に、反射壁境界の代わりに基本復帰 (elementary return) 境界を用いる方法である [7, 8]。ポアソン到着の場合には、システムが空になっている時間期間の長さ  $T_0$  は明らかに指数分布にしたがっている。その期間の後、系内客数過程の見本関数は  $Q(\cdot) = 1$  から再び開始される。基本復帰境界は、このように指数分布に

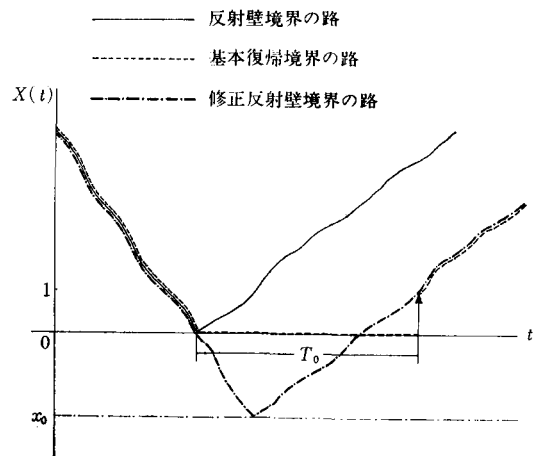


図 2 拡散過程の境界条件

したがう境界上での滞在時間を持ち、その後、領域の内側への跳躍を行なう確率過程の境界条件であり、軽負荷時の特性量の挙動を適切に表現しているといえる。この境界条件には指数分布の無記憶性が本質的に効いているため、到着分布が一般への拡張はおそらく困難であると考えられる。一般到着分布の場合にも適用できる発見的な方法として、反射壁境界を原点での見本関数の滞在時間に応じて負の領域へ移動させる方法が知られている [15, 16, 2]。この際問題となるのは反射壁境界の移動幅  $x_0$  の決め方で、原点における確率質量が未知なために、これまではっきりとはわかっていなかった。筆者は拡散過程の初通過時間 (first passage time) に着目して、単一窓口待ち行列に対しては、反射壁境界の位置を、

$$x_0 = \min \left( 0, \frac{a}{2b} \log \left\{ \frac{2b}{a} \cdot \frac{1 - bE[T_0]}{1 - \exp(-2b/a)} \right\} \right) \quad (11)$$

に移動させればよいことを示した [17, 第 3 章]。 (ただし、 $a = a(1) = \lambda - \mu$ ,  $b = b(1) = \lambda^3 \sigma_a^2 + \mu^3 \sigma_s^2$ .) しかし、(11) 式中の  $E[T_0]$  は、到着分布が指数分布の時を除いて陽な表現が未だに導かれていない。したがって、実際の計算にあたっては、 $T_0$  を何か適当な確率変数で近似する必要がある。こういった場合よく用いられるのが到着時間間隔の定常残余寿命 (stationary residual life time) で、この近似の下では、

$$E[T_0] = \frac{\lambda^2 \sigma_a^2 + 1}{2\lambda} \quad (12)$$

を(11)に代入すればよい。この反射壁境界を移動させる方法も、拡散パラメータが状態に依存する場合には解決すべき問題点が残っている。

### 3.3 密度関数の離散化

拡散パラメータと境界条件が決定されれば、拡散過程はほぼ一意に決定できる。つまり密度関数  $p(x, t | x_0)$  が計算できるわけだが、この密度関数の解釈の仕方にまだいくらかの自由度が残されている。それは元の確率過程がもっていた離散性を近似解が失っていることに起因しており、連続近似である拡散近似にとっては避けて通れない問題である。特に、系内客数過程の拡散近似の場合には、密度関数を離散化して系内客数の離散分布を計算する必要が生ずる。 $\pi_n$  を定常状態において系内客数が  $n$  人である確率とし、 $p(x)$  を対応する密度関数であるとする、離散化は具体的には次のように行なわれる。

$$\pi_n = \int_n^{n+1} p(x) dx \quad n=0, 1, \dots \quad (13)$$

あるいは、

$$\pi_n = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} p(x) dx \quad n=1, 2, \dots \quad (14)$$

離散分布  $\{\pi_n\}$  が(14)式で定義される時には、定常確率  $\pi_0$  を与えてやる必要があり、さらにその後、正規化の操作も必要になってくる。また、(13)式で定義される時にも、なんらかの補正が必要であることが報告されている[18]。このように密度関数の離散化には未解決の部分が多く、この処理を誤ると近似解の精度を低下させることにもなりかねない。同様の問題は、特性量のモーメントの定義にも見出される。

## 4. 近似の有用性

拡散近似はおおよそ以上の段階を経て行なわれるわけだが、それは如何なる有用性をもっているのだろうか？ 近似である以上、あらゆる場合に正確であることは望めないが、精度の点は別にし

て、拡散近似は次の3つの大きな特徴をもっていると考えられる。

- (i) 厳密解が得られていない複雑な待ち行列の解析が可能なこと。
- (ii) 到着やサービス等に関する分布の完全な情報を必要としないこと。すなわち、近似解がロバスト(robust)であること。
- (iii) 近似解が陽な形で得られること。

特徴(i)は、ネットワーク型の待ち行列(queueing networks)の解析に典型的に見ることができ、Jackson[14]によって導入された開放型待ち行列ネットワーク(open queueing networks)を一般化したシステムや、古典的な修理事務問題の一般化と考えられる閉鎖型待ち行列ネットワーク(closed queueing networks)システム等が、拡散近似によって解かれている[18, 19, 21]。その他、計算機システムの性能評価の問題に対しても、拡散近似は多くの成果をあげている(たとえば[4, 5, 6, 7])。特徴(ii)については、重負荷極限定理で示されているように、到着やサービスの分布の最初の2つのモーメントだけが近似解に効いていて、その他の詳しい情報を必要としないことから明らかである。このことは、分布全体を推定することがむずかしいことから、実際に近似解を利用する際には非常に重要な特徴である。もし分布形がわかっているときには、そのLaplace-Stieltjes変換を用いて近似解を補正する方法も知られており[6]、3次以上のモーメントの微妙な違いを取り入れることも可能である。特徴(iii)は、すべての拡散近似解がもち合わせているわけではなく、特性量の種類や拡散パラメータの決め方等によっては成り立たない場合がある。たとえば、GI/G/s待ち行列の系内客数過程の拡散近似で、拡散パラメータを(9)式で定義すると、定常確率  $\{\pi_n\}$  は陽な形で求めることができない。また、M/G/1待ち行列の稼働期間の拡散近似では、いくつかの簡単なサービス分布を除いて、稼働期間分布を積分形でしか求めることができないことがわかっている。

る[11]. しかし, 大半の拡散近似解は陽な形で与えられており, プログラム電卓程度の計算機で容易にその値を計算することができる.

例 以上の特徴を如実に示す例として,  $M/G/s$  待ち行列の待ち時間分布に対する拡散近似解を与えておく.  $W$  を定常状態における待ち時間を表わす確率変数とすると,  $\rho < 1$  のとき,

$$P\{W \leq t\} = \begin{cases} 0, & t < 0 \\ 1 - \Pi e^{-\gamma t}, & t \geq 0 \end{cases} \quad (15)$$

という簡約な形で求めることができる [17, 第 6 章]. ただし,

$$\gamma = \lambda \left( \exp \left\{ -\frac{2(\lambda - s\mu)}{\lambda^2 \sigma_a^2 + s\mu^2 \sigma_s^2} \right\} - 1 \right) > 0 \quad (16)$$

であり,  $\Pi = P\{W > 0\}$  も到着やサービス分布の平均, 分散だけからなる陽な式で表わせる.

## 5. むすび

待ち行列の近似モデルの 1 つである拡散近似についてその考え方と有用性を述べたが, 発想の単純さとは裏腹に, 対象となるシステムの特異性を生かした細かい工夫がいかにか大切かをご理解いただけたと思う. 天使のような大胆さと悪魔のような繊細さ(!?), この 2 つの相異なる性質を合わせもたせることが, 拡散近似に限らず, すべての近似において重要なことではないだろうか. 使いやすい正確な近似解も, こういった近似の試みの中から生まれてくると思われる.

### 参考文献

- [1] Billingsley, P., *Convergence of Probability Measures*, Wiley, New York, 1968.
- [2] Biswas, S. K., and Sunaga, T., "Diffusion Approximation Method for Multi-Server Queueing System with Balking," *Journal of the Operations Research Society of Japan*, **23**, 368-386 (1980).
- [3] Feller, W., "Diffusion Processes in One Dimension," *Transactions of the American Mathematical Society*, **77**, 1-31 (1954).
- [4] Gaver, D. P., "Analysis of Remote Terminal Backlogs under Heavy Demand Conditions," *Journal of the Association for Computing Machinery*, **18**, 405-415 (1971).
- [5] \_\_\_\_\_, and Schedler, G. S., "Processor Utilization in Multiprogramming Systems via Diffusion Approximation," *Operations Research*, **21**, 569-576 (1973).
- [6] \_\_\_\_\_, and \_\_\_\_\_, "Approximate Models for Processor Utilization in Multiprogrammed Computer Systems," *SIAM Journal on Computing*, **2**, 183-192 (1973).
- [7] Gelenbe, E., "On Approximate Computer System Models," *Journal of the Association for Computing Machinery*, **22**, 261-269 (1975).
- [8] \_\_\_\_\_, and Pujolle, G., "The Behaviour of a Single Queue in a General Queueing Network," *Acta Informatica*, **7**, 123-136 (1976).
- [9] Goel, N. S., and Dyn, N. R., *Stochastic Models in Biology*, Academic Press, New York, 1974. (寺本 他訳: 『生物学における確率過程の理論』 数理解析とその周辺 22, 産業図書, 1978.)
- [10] Halachmi, B., and Franta, W. R., "A Diffusion Approximation to the Multi-Server Queue," *Management Science*, **24**, 522-529 (1978).
- [11] Heyman, D. P., "An Approximation for the Busy Period of the  $M/G/1$  Queue Using a Diffusion Model," *Journal of Applied Probability*, **11**, 159-169 (1974).
- [12] \_\_\_\_\_, "A Diffusion Model Approximation for the  $GI/G/1$  Queue in Heavy Traffic," *The Bell System Technical Journal*, **54**, 1637-1646 (1975).
- [13] Iglehart, D. L., and Whitt, W., "Multiple Channel Queues in Heavy Traffic. I," *Advances in Applied Probability*, **2**, 150-177 (1970).
- [14] Jackson, J. R., "Networks of Waiting Lines," *Operations Research*, **5**, 518-521 (1957).
- [15] Kimura, T., Ohno, K., and Mine, H., "Diffusion Approximation for  $GI/G/1$  Queue-

- ing Systems with Finite Capacity : I-The First Overflow Time," *Journal of the Operations Research Society of Japan*, **22**, 41-68 (1979).
- [16] \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_, "Diffusion Approximation for GI/G/1 Queueing Systems with Finite Capacity : II-The Stationary Behaviour," *ibid.*, **22**, 301-320 (1979).
- [17] Kimura, T., *Studies on Diffusion Approximation for Queueing Systems*, Doctoral Dissertation, Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, December 1980.
- [18] Kobayashi, H., "Application of the Diffusion Approximation to Queueing Networks I : Equilibrium Queue Distributions," *Journal of the Association for Computing Machinery*, **21**, 316-328 (1974).
- [19] \_\_\_\_\_, "Application of the Diffusion Approximation to Queueing Networks II : Nonequilibrium Distributions and Applications to Computer Modeling," *ibid.*, **21**, 459-469 (1974).
- [20] Колмогоров, А. Н., Основные Понятия Теории Вероятностей, издание 2-е, Наука, Москва, 1974. (根本訳:『確率論の基礎概念』第二版, 東京図書, 1975.)
- [21] Lemoine, A. J., "Networks of Queues-A Survey of Weak Convergence Results," *Management Science*, **24**, 1175-1193 (1978).
- [22] Newell, G. F., *Applications of Queueing Theory*, Chapman & Hall, London, 1971. (森村, 森訳:『待ち行列理論の応用—その新しい方法』サイエンス社, 1973.)
- [23] 逆瀬川浩孝, "待ち行列における近似モデル", オペレーションズ・リサーチ, **25**, 794-800 (1980).
- [24] Sunaga, T., Kondo, E., and Biswas, S. K., "An Approximation Method Using Continuous Models for Queueing Problems," *Journal of the Operations Research Society of Japan*, **21**, 29-44 (1978).