

シリーズ 情報科学における確率モデル **10**

Series on Stochastic Models in Informatics and Data Science

# 最良選択問題の諸相

——秘書問題とその周辺——

玉置 光司【著】

コロナ社



シリーズ 情報科学における確率モデル  
編集委員会

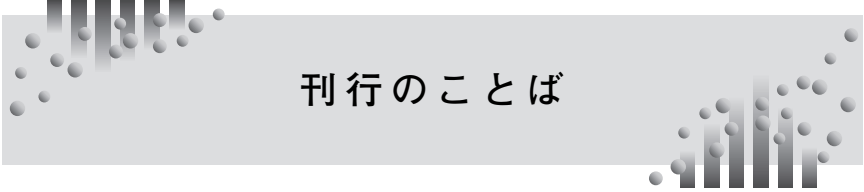
編集委員長

博士（工学） 土肥 正（広島大学）

編集委員

博士（工学） 栗田多喜夫（広島大学）

博士（工学） 岡村 寛之（広島大学）



## 刊行のこ と ば

われわれを取り巻く環境は、多くの場合、確定的というよりもむしろ不確実性にさらされており、自然科学、人文・社会科学、工学のあらゆる領域において不確実な現象を定量的に取り扱う必然性が生じる。「確率モデル」とは不確実な現象を数理的に記述する手段であり、古くから多くの領域において独自のモデルが考案されてきた経緯がある。情報化社会の成熟期である現在、幅広い裾野をもつ情報科学における多様な分野においてさえも、不確実性下での現象を数理的に記述し、データに基づいた定量的分析を行う必要性が増している。

一言で「確率モデル」といっても、その本質的な意味や粒度は各個別領域ごとに異なっている。統計物理学や数理生物学で現れる確率モデルでは、物理的な現象や実験的観測結果を数理的に記述する過程において不確実性を考慮し、さまざまな現象を説明するための描写をより精緻化することを目指している。一方、統計学やデータサイエンスの文脈で出現する確率モデルは、データ分析技術における数理的な仮定や確率分布関数そのものを表すことが多い。社会科学や工学の領域では、あらかじめモデルの抽象度を規定したうえで、人工物としてのシステムやそれによって派生する複雑な現象をモデルによって表現し、モデルの制御や評価を通じて現実に役立つ知見を導くことが目的となる。

昨今注目を集めている、ビッグデータ解析や人工知能開発の核となる機械学習の分野においても、確率モデルの重要性は十分に認識されていることは周知の通りである。一見して、機械学習技術は、深層学習、強化学習、サポートベクターマシンといったアルゴリズムの違いに基づいた縦串の分類と、自然言語処理、音声・画像認識、ロボット制御などの応用領域の違いによる横串の分類によって特徴づけられる。しかしながら、現実の問題を「モデリング」するためには経験とセンスが必要であるため、既存の手法やアルゴリズムをそのまま

## ii 刊 行 の こ と ば


適用するだけでは不十分であることが多い。

本シリーズでは、情報科学分野で必要とされる確率・統計技法に焦点を当て、個別分野ごとに発展してきた確率モデルに関する理論的成果をオムニバス形式で俯瞰することを目指す。各分野固有の理論的な背景を深く理解しながらも、理論展開の主役はあくまでモデリングとアルゴリズムであり、確率論、統計学、最適化理論、学習理論がコア技術に相当する。このように「確率モデル」にスポットライトを当てながら、情報科学の広範な領域を深く概観するシリーズは多く見当たらず、データサイエンス、情報工学、オペレーションズ・リサーチなどの各領域に点在していた成果をモデリングの観点からあらためて整理した内容となっている。

本シリーズを構成する各書目は、おのおのの分野の第一線で活躍する研究者に執筆をお願いしており、初学者を対象とした教科書というよりも、各分野の体系を網羅的に著した専門書の色彩が強い。よって、基本的な数理的技法をマスターしたうえで、各分野における研究の最先端に上り詰めようとする意欲のある研究者や大学院生を読者として想定している。本シリーズの中に、読者の皆さんのアイデアやイマジネーションを掻き立てるような座右の書が含まれていたならば、編者にとっては存外の喜びである。

2018年11月

編集委員長 土肥 正



# まえがき

期待利益の最大化あるいは期待損失の最小化というような明確な最適化基準の下で、対象とする確率過程の変化を逐次観測しながらどの時点で所与の行動をとるかを決定するのが確率最適化 (stochastic optimization) と呼ばれる問題である。その中でも、どのタイミングで観測を停止する (多くの場合、そのときの対象を選択して停止する) かを決める問題が最適停止問題 (optimal stopping problem) であり、秘書問題 (secretary problem) は秘書のポストを求めて出現する応募者の中から望ましい人を採用するという状況設定を用いた特殊な最適停止問題と考えられる。最適停止問題を解くことは一般に困難であるが、秘書問題はその特殊な構造の見返りとして実際に解くことができる問題も多く、そのエレガントな解が人々を引き付けてきた。

秘書問題の基本的な枠組みはつぎのように述べられる。秘書 1 人の採用に対して  $n$  人の応募があり、採用者は応募者の能力を 1 (最良) から  $n$  (最悪) まで順位 (絶対順位) によって評価することができる。したがって、全員を一堂に集めて集団面接すれば容易にその中の最良を選ぶことができるが、応募者が毎時 1 人ずつランダムな順序で面接に臨む場合はどうなるか。この場合、採用者の利用可能な情報は応募者の相対順位 (すでに面接をすませた人の中の順位) だけで、それに基づいて毎時採用か不採用 (パス) かを決めなければならない。一度パスしたら後からは採用できない。また、最初の  $n - 1$  人をパスしたら最後の応募者を採用しなければならない。

最適化基準に関して二つのモデルを紹介しよう。最もよく知られたモデルはベスト (最良) をねらう問題である。この問題は自然対数の底  $e$  の逆数  $e^{-1}$  ( $\approx 0.368\dots$ ) と関係している。すなわち、 $n$  が大きくなると、全体の約 37% ( $\approx e^{-1}n$ ) の応募者をパスし、その後に出現する最初の相対順位 1 の応募者を採用

することが最適となる。また、この最適ルールの下で実際にベストが選ばれる確率も  $e^{-1}$  に近づく (定理 1.1)。Lindley<sup>40)</sup>†は、このモデルを最良の伴侶を選ぶ結婚問題に適用し、つぎのような具体例を挙げている。18 歳から 40 歳までを適齢期と考える男性の最適選択ルールは、 $18 + (40 - 18) \times 0.368 \approx 26$  となることから、26 歳までは観察するにとどめ、その後最初の相対順位 1 の女性にプロポーズせよというものである。 $n = 40 - 18 = 22$  とみなし、これを大きな数と考えているようであるが、実際の出現数は 22 人とはかぎらない。このように出現数に不確実性がある場合は、むしろ 6 章で紹介する Bruss<sup>4)</sup> の連続時間モデルのほうが適切と考えられる (定理 6.7, 系 6.1)。両者は似た結果を与えるが、Bruss モデルは出現数の不確実性に対して頑健性がある。

もう一つのモデルは選ばれた応募者の順位 (の期待値) を最小にする問題である。 $n$  人の中の最良が順位 1, そのつぎが順位 2, ..., 最悪が順位  $n$  であったから、これも適切な最適化基準と考えられる。このとき  $n$  が大きくなると、最適ルールの下で達成される期待順位はつぎの無限積

$$\prod_{j=1}^{\infty} \left( \frac{j+2}{j} \right)^{1/(j+1)} = \left( \frac{3}{1} \right)^{1/2} \left( \frac{4}{2} \right)^{1/3} \left( \frac{5}{3} \right)^{1/4} \cdots \approx 3.8695$$

に近づく。なんと、平均的に順位 4 以下の人を選ぶことができる (定理 1.2)。最適ルールは少し複雑である。まず、全体の 26% の応募者をパスする。その後、相対順位 1 の応募者が出現すればただちに採用するが、もし 45% まで面接をつづけても該当者が出現しなければ、その後は相対順位 1 でも 2 でも早く出たほうを採用する。さらに、56% まで面接をつづけても該当者が出現しなければ、相対順位 3 以下まで基準を緩める。このように、最適ルールは時間の進行と共に採用基準を緩和する。

最初のモデルは最良選択 (best choice, BC) 問題と呼ばれ、2 番目のモデルは順位最小化 (rank minimization, RM) 問題と呼ばれる。最適化基準に従った分類である。利用可能な情報に従った分類も可能である。上に述べた二つのモデルでは、利用可能な情報は応募者の相対順位だけであった。このような問

† 肩付き数字は、巻末の引用・参考文献の番号を表す。

題を無情報型 (no information, NI) と呼ぶ。この対極にあるのが完全情報型 (full information, FI) と呼ばれる問題で、そこでは  $k$  番目の応募者に付随する評価値  $X_k$ ,  $1 \leq k \leq n$  が観測され、それが利用可能な情報となる。通常,  $X_1, \dots, X_n$  は区間  $(0, 1)$  上で一様に分布する独立確率変数列と仮定される。完全情報型は無情報型より情報が多いので、採用者にとって (期待利得が大きい) ベターな決定が期待できる。

本書の構成は以下のようになっている。1 章では、最適化基準と利用可能な情報の組合せからなる四つの問題、すなわち無情報型最良選択問題、無情報型順位最小化問題、完全情報型最良選択問題、完全情報型順位最小化問題を紹介する。本書では、これらはしばしば英語の頭文字をとって、NIBC, NIRM, FIBC, FIRM と略記される。2 章では、NIBC の多方面への一般化を試みる。3 章では、NIRM に関係した変形モデルをいくつか紹介する。4 章の Sum-the-odds 定理も、NIBC の一般化と考えられるが、1-sla (1-stage look-ahead) ルールとの関係から興味深い応用につながる。5 章の問題は Ferguson<sup>21)</sup> が提起した問題で、秘書問題のルーツといえる数当てゲームのゲーゴル (Googol) と深く関係している (ゲーゴルについては、そこで改めて紹介する)。ここまでは応募者総数  $n$  を既知と仮定しているが、6 章では未知の場合への拡張を試みる。7 章では、期間最大化という新しい最適化基準の下で秘書問題を考える。期間最大化問題と最良選択問題の間の興味深い対応関係も示される。秘書問題では、 $n$  を大きくしたときの特性値の挙動に大きな関心が寄せられるが、これを調べることは、特に完全情報型問題の場合は容易ではない。8 章では、この困難を克服する試みとして提案された PPP (planar Poisson process) によるアプローチを紹介する。これは、完全情報型の極限問題を平面ポアソン過程の上で考えようというものである。

秘書問題の歴史については Ferguson<sup>21)</sup> に詳しい。それによれば、活字になった最初の秘書問題は Scientific American に掲載された Gardner<sup>27)</sup> のゲーゴルに関する記事と考えられている。したがって、秘書問題は 60 年を超す歴史を有する。1960 年代に入ると、Lindley<sup>40)</sup>, Dynkin<sup>19)</sup>, Chow, Moriguti, Robbins

and Samuels<sup>13)</sup>, Gilbert and Mosteller<sup>30)</sup>, Dynkin and Yushkevitch<sup>20)</sup>など著名な研究者による重要な研究がつぎつぎと発表された。中でも特筆すべきは Gilbert and Mosteller<sup>30)</sup> で、この論文に含まれる多くのモデルはその後の発展の萌芽となり、いまなお、この分野を目指す人の必読論文でありつづけている。少し遅れるが、Chow, Robbins and Siegmund<sup>14)</sup>, DeGroot<sup>16)</sup> も最適停止の研究に大きな影響を与えた。

秘書問題のサーベイ論文としては Samuels<sup>81)</sup>, Ferguson<sup>21)</sup>, Freeman<sup>26)</sup> が推奨される。坂口<sup>75)</sup>, 玉置<sup>93)</sup> からは文献に関して若干の追加情報が得られる。研究集会の報告書としては、アメリカ数学会の Contemporary Mathematics 125 (1992) やポーランドの数学会誌 *Matematyka Stosowana* (Mathematica Applicanda 44(1) (2016)) が参考になる。秘書問題に特化した本としては、Berezovskiy and Gnedin<sup>3)</sup> のロシア語本が知られている。Ferguson<sup>22)</sup> の電子テキスト “Optimal Stopping and Applications” は、広く最適停止問題を研究する人にとって格好のテキストである。この本は秘書問題にもスペースを割いていて、本書の執筆に際しても利用させていただいた。和書では最近、高木英明著「秘書問題入門」<sup>88)</sup> が出版された。豊富な計算例を通して秘書問題が体感できる。穴太克則著「タイミングの数理」<sup>1)</sup> も参考になる。

本書は厳密な理論展開というよりは直感的理解を重んじた記述になっている。したがって、読む上で高度な数学は必要とされない。理系学部で学ぶ微分積分と応用確率論の知識があれば十分である。ただし、結果の紹介にとどまる箇所が一部ある。本書を通して、多くの人に秘書問題の面白さを知っていただければ幸いである。

最後に、研究生活でお世話になった方々に謝意を表したい。筆者をこの分野に導いて下さった坂口実教授（大阪大学）に感謝いたします。前述の論文をご覧になれば、先生の秘書問題への強い思いが伝わってきます。Thomas Ferguson (UCLA), Thomas Bruss (Université Libre de Bruxelles), Krzysztof Szajowski (Wrocław University of Technology), Vladimir Mazalov (Karelian Research Center of the Russian Academy of Science) の諸先生方の 30 年を




超えるご交誼こうぎに感謝いたします。この中の3氏とは共同研究の機会にも恵まれました。生前親しくご指導いただいた大野勝久先生（名古屋工業大学）の薫陶くんとうも忘れられません。しばしば有益なご助言をいただいた中井達先生（千葉大学）のご厚情にも感謝いたします。数値計算などでお世話になった畏友いゆう、太田拓男、浅倉史興，玉井清二，有澤健治の諸氏に改めてお礼申し上げます。ITに精通した王 琦先生（長崎総合科学大学）の献身的支援に深く感謝いたします。遅筆の筆者に辛抱強くお付き合いいただき，原稿のチェックと適切なご助言をいただいたコロナ社の皆様に厚くお礼申し上げます。最後に本書の執筆をおすすめいただいた編集委員長の土肥正教授（広島大学大学院）のご厚意と貴重なコメントに深甚なる謝意を表します。

本書を父母に捧げる。

2023年5月

玉置 光司



# 目 次

## 第 1 章 秘書問題の主要モデル

---

1.1 秘書問題	1
1.2 無情報型モデル	3
1.2.1 無情報型最良選択問題	6
1.2.2 無情報型順位最小化問題	10
1.3 完全情報型モデル	16
1.3.1 完全情報型最良選択問題	16
1.3.2 完全情報型順位最小化問題	24

## 第 2 章 無情報型最良選択問題の展開

---

2.1 拒否とリコール	32
2.1.1 Petruccelli モデル	32
2.1.2 もう一つの拒否モデル	40
2.2 候補者選択問題	46
2.2.1 割引を考慮した NIBC	48
2.2.2 坂口モデル	52
2.2.3 1-sla ルールの最適性	54
2.3 利得の一般化	58
2.3.1 ベストあるいはセカンドベストの選択	60
2.3.2 セカンドベストの選択	65
2.4 トレーニングサンプル付最良選択問題	68
2.4.1 トレーニングサンプル	68
2.4.2 漸近挙動	72

### 第3章 無情報型順位最小化問題の展開

3.1	メモリの制限	79
3.1.1	メモリサイズ	80
3.1.2	メモリサイズ1のNIRM	82
3.1.3	無限問題	85
3.2	NIRMの簡易ルール	91
3.2.1	短縮型ルール	91
3.2.2	漸近挙動とNHPP	95

### 第4章 Sum-the-odds 定理とその展開

4.1	Sum-the-odds 定理	106
4.2	Sum-the-odds 定理の一般化	114
4.2.1	独立でないベルヌーイ試行	115
4.2.2	FIBCの一般化	120
4.3	Sum-the-multiplicative-odds 定理	123

### 第5章 Fergusonの秘書問題

5.1	グーゴル	128
5.2	Fergusonの生成ルール	130
5.3	Gnedinの生成ルール	135

### 第6章 出現数が未知の場合の最良選択問題

6.1	不確実性の導入	142
-----	---------	-----

6.2 無情報型問題 .....	143
6.2.1 Presman and Sonin モデル	143
6.2.2 Samuel-Cahn モデル	157
6.2.3 Bruss の連続時間モデル	160
6.3 完全情報型問題 .....	163
6.3.1 Porosiński モデル	164
6.3.2 Samuel-Cahn モデル	175
6.4 Petruccelli の部分情報型最良選択問題 .....	180
6.4.1 PET	180
6.4.2 PET と POR の奇妙な一致	183

## 第7章 期間問題

---

7.1 応募者数が既知の期間問題 .....	191
7.1.1 無情報型期間問題	191
7.1.2 完全情報型期間問題	195
7.2 応募者数が未知の無情報型期間問題 .....	199
7.3 最良選択問題と期間問題の交互対応 .....	206

## 第8章 PPP と FI モデル

---

8.1 PPP .....	214
8.2 FIBC .....	217
8.2.1 最適ルール	217
8.2.2 成功率	217
8.3 PET と FIBC .....	225
8.3.1 最適ルール	226
8.3.2 成功率	229
8.4 POR と FIBC .....	234
8.4.1 最適ルール	234

付 録	241
A.1 マルチンゲール停止定理	241
A.2 単調ルールの下での期待利得	242
引用・参考文献	245
あ と が き	252
索 引	254

# 1

## 秘書問題の主要モデル

本章では秘書問題の四つの主要モデルを紹介する。2種類の情報構造と2種類の最適化基準の組合せである。

### 1.1 秘書問題

秘書問題 (secretary problem) は、秘書のポストを求めて出現する応募者の中から望ましい人を採用する (採用を停止と呼ぶことがある) という状況設定を用いた特殊な最適停止問題 (optimal stopping problem) である。その中でもよく知られた古典的秘書問題 (classical secretary problem) は、Ferguson<sup>21)</sup> になれば、その構成要素は以下のように列挙される。

1. 秘書の採用は1人である。
2. 応募者総数  $n$  は既知である。便宜上、毎時1人ずつ出現するものとする。
3. 採用者 (意思決定者) は応募者を (面接によって) 評価することができる。すなわち、応募者を一堂に集めれば、1 (最良) から  $n$  (最悪) まで順位 (絶対順位 (absolute rank)) を付けることができる。実際には1人ずつ出現するので、採否の決定は応募者の相対順位 (relative rank) (すでに面接をすませた人の中での順位) に基づいて行われる。
4. 応募者の出現順序はランダムである。すなわち、 $n!$  通りの順列 (出現順序) は等確率で起こる。
5. 応募者を面接したら、採用か不採用かをただちに決める。前者の場合は

## 2 1. 秘書問題の主要モデル

その時点で採用プロセスは終了する（停止する）が、後者の場合は（もし存在すれば）つぎの応募者との面接に進む。

6. 一度断った人を後から採用することは許されない。
7. 採用者の目標は最良（絶対順位 1）の応募者を選ぶことである。すなわち、最良を選べば利得は 1 であるが、それ以外を選べば利得は 0 である。

秘書問題の発展とは、上記各項目の変更ないしは拡張と考えられる。例えば、項目 3. は応募者の評価のために（採用者にとって）利用可能な情報が相対順位だけであることを述べているが、具体的な数値で評価できる場合は、3. をつぎの 3'. で置き換えた問題がよく研究されている（Gilbert and Mosteller<sup>30</sup>，Sakaguchi<sup>73</sup>）。

- 3'.  $k$  番目の応募者の評価値を  $X_k$ ， $1 \leq k \leq n$  で表す。 $X_1, \dots, X_n$  はたがいに独立で同一の分布に従う連続確率変数列で、それらは共通の既知の分布関数  $F(x)$  をもつ。採否の決定は、それまでに観測された評価値に基づいて行われる。

連続性の仮定は、異なる応募者の評価値が等しくなる（タイの）可能性を排除するためである。利得が順位に依存する場合、 $F(x)$  を区間  $(0, 1)$  上の一様分布関数と仮定しても一般性を失わない。なぜならば、 $Y_k = F(X_k)$  と変換すれば、よく知られているように  $Y_k$  は区間  $(0, 1)$  上の一様確率変数となり、この変換によって  $X_1, \dots, X_n$  の間の大小関係は  $Y_1, \dots, Y_n$  の間でも保たれるからである。3'. においては、利用可能な情報を相対順位に落とし込むことができるので、3. の場合より情報が豊富である。この観点から、3. の場合を無情報型 (no information) と呼び、3'. の場合を完全情報型 (full information) と呼ぶ。前者を NI，後者を FI と略記することがある。もう一つの変更は最適化基準に関するもので、項目 7. をつぎの 7'. で置き換えた問題である（Lindley<sup>40</sup>，Chow，Moriguti, Robbins and Samuels<sup>13</sup>，Robbins<sup>65</sup>）。

- 7'. 採用者の目標は選んだ応募者の期待順位を最小にすることである。

項目 7. はベスト（最良をベストと呼ぶことがある）以外は価値がないという

極端な基準であるが、応募者は順位が小さいほどベターであるから、項目  $7'$  は自然で妥当な基準と考えられる。最適化基準の観点から、 $7$  の場合を**最良選択問題** (best choice problem) と呼び、 $7'$  の場合を**順位最小化問題** (rank minimization problem) と呼ぶ。前者を BC、後者を RM と略記することができる。この分類に従えば、古典的秘書問題は無情報型最良選択問題、すなわち NIBC と見ることができる。本章では  $3$  と  $3'$ 、 $7$  と  $7'$  の組合せからなる以下の四つの問題を紹介する。

- 無情報型最良選択問題 (NIBC)。
- 無情報型順位最小化問題 (NIRM)。
- 完全情報型最良選択問題 (FIBC)。
- 完全情報型順位最小化問題 (FIRM)。

## 1.2 無情報型モデル

無情報型モデル (NI モデル) の特徴はすでに述べたが、改めて簡単に説明しよう。秘書のポストを求めて  $n$  人の応募者がランダムに出現する。採用者はその都度、面接を行い、観測された相対順位に基づいて、その応募者の採否を決める。一度断った人を後から採用することは許されない。最初の  $n-1$  人を断ったら最後 ( $n$  番目) の応募者は必ず採用しなければならない (本書では、特に断らないかぎり、最後の応募者の採用を仮定する)。採用者の目的は与えられた問題の最適 (採用) ルールを求め、可能であればそのルールの下で達成される (期待) 利得を計算することである。 $n$  を大きくしたときの**漸近挙動** (asymptotic behavior) にも関心がある。

本書では、しばしば  $k$  番目の応募者 (時刻  $k$  に出現する応募者) の相対順位を  $R_k$  で、絶対順位を  $A_k$  で表す。したがって、 $A_k$  が集合  $\{A_1, \dots, A_k\}$  の中で  $j$  番目に小さいとき、 $R_k = j$  となる ( $1 \leq j \leq k \leq n$ )。最初に、NI モデルの重要な性質を二つ与える。



**補題 1.1** 確率変数  $R_1, \dots, R_n$  はたがいに独立で、 $R_k$  の分布は一様である。すなわち

$$P\{R_k = j\} = \frac{1}{k}, \quad 1 \leq j \leq k, 1 \leq k \leq n \quad (1.1)$$

で与えられる。

**【証明】**  $1, \dots, n$  の  $n$  個の要素からなる  $n!$  通りの順列の集合を  $Q$  とする。ランダム性の仮定（前述の項目 4.）より、任意の順列  $(a_1, \dots, a_n) \in Q$  に対して

$$P\{A_1 = a_1, \dots, A_n = a_n\} = \frac{1}{n!} \quad (1.2)$$

である。 $R_1, \dots, R_n$  の実現値は  $A_1, \dots, A_n$  を一意に決めることに注意する。例えば、 $n = 4$  で  $R_1 = 1, R_2 = 1, R_3 = 3, R_4 = 2$  のとき、逆向きにたどれば、 $A_4 = 2, A_3 = 4, A_2 = 1, A_1 = 3$  であることがわかる。したがって、 $R_1, \dots, R_n$  の生起確率是对應する  $A_1, \dots, A_n$  の生起確率に等しい。これは式 (1.2) より任意の  $r_j = 1, \dots, j, 1 \leq j \leq n$  に対して

$$P\{R_1 = r_1, \dots, R_n = r_n\} = \frac{1}{n!} \quad (1.3)$$

を意味する。この同時分布から、 $R_n$  の周辺分布はつぎのように計算される。

$$\begin{aligned} P\{R_n = r_n\} &= \sum_{r_1} \cdots \sum_{r_{n-1}} P\{R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = r_n\} \\ &= \sum_{r_1} \cdots \sum_{r_{n-1}} \frac{1}{n!} \\ &= \frac{(n-1)!}{n!} \\ &= \frac{1}{n} \end{aligned}$$

3 番目の等式は  $r_j$  が  $j$  通りの値をとるので、 $n-1$  重和の項数が全部で  $1 \cdot 2 \cdots (n-1) = (n-1)!$  個あることから得られる。同様にして、 $R_k$  の周辺分布が  $P\{R_k = r_k\} = 1/k$ , すなわち、式 (1.1) が示される。このことは、式 (1.3) との比較から

$$P\{R_1 = r_1, \dots, R_n = r_n\} = P\{R_1 = r_1\} \cdots P\{R_n = r_n\}$$

の成立を意味し、 $R_1, \dots, R_n$  の独立性が示される。

◇

**補題 1.2**  $k$  番目の応募者の相対順位が  $j$  であるとき、以下が成立する。

(a) この応募者の順位が  $i$  である確率は次式で与えられる。

$$P\{A_k = i \mid R_k = j\} = \frac{\binom{i-1}{j-1} \binom{n-i}{k-j}}{\binom{n}{k}} \quad (1.4)$$

ただし,  $j \leq i \leq n - k + j$ ,  $1 \leq k \leq n$ 。

(b) この応募者の期待順位は次式で与えられる。

$$E[A_k \mid R_k = j] = \left(\frac{n+1}{k+1}\right)j \quad (1.5)$$

**【証明】**

(a)  $S$  を絶対順位  $1, \dots, i-1$  の応募者の集合,  $T$  を絶対順位  $i+1, \dots, n$  の応募者の集合, そして  $U$  を絶対順位  $i$  (だけ) の応募者の集合とする。求める確率は,  $n$  人から  $k$  人をランダムに選ぶとき,  $S$  から  $j-1$  人選び,  $T$  から  $k-j$  人選び,  $U$  から 1 人 (絶対順位  $i$ ) を選ぶ確率であるからただちに式 (1.4) を得る。


(b) 最も簡単な証明は条件付確率の和が 1 となることを利用することである。すなわち,

$$\sum_{i=j}^{n-k+j} P\{A_k = i \mid R_k = j\} = 1 \text{ より}$$

$$\sum_{i=j}^{n-k+j} \binom{i-1}{j-1} \binom{n-i}{k-j} = \binom{n}{k}, \quad j \leq k \leq n \quad (1.6)$$

を得る。したがって

$$\begin{aligned} E[A_k \mid R_k = j] &= \sum_{i=j}^{n-k+j} iP\{A_k = i \mid R_k = j\} \\ &= \sum_{i=j}^{n-k+j} i \frac{\binom{i-1}{j-1} \binom{n-i}{k-j}}{\binom{n}{k}} \end{aligned}$$



## あ と が き

本書の副題は「秘書問題とその周辺」であった。「まえがき」で秘書問題が特殊な最適停止問題であることは述べたが、それではその境界はどこにあるのだろうか。従来、研究者は思い思いのイメージで秘書問題という言葉を使ってきたが、この言葉の定義を試みたのが Ferguson<sup>21)</sup> である。つぎのような定義を与えた。

A secretary problem is a sequential observation and selection problem in which the payoff depends on the observations only through their relative ranks and not otherwise on their actual values.

この定義に従えば、これまで本書で扱った問題は必ずしもすべてが秘書問題ではなかったかもしれない。例えば、Sum-the-odds 定理を秘書問題に含めることには異論が出るかもしれない。副題を付けたのはこのような理由からである。

紙数の関係で本書では紹介できなかったが、Ferguson の定義に照らせば秘書問題に入る可能性がある二つの問題を簡単に紹介して、本書を終える。

- (a) **Tamaki and Shanthikumar**<sup>101)</sup> :  $n$  個の評価値  $X_k$ ,  $1 \leq k \leq n$  が独立で同一の分布に従うとき、FIBC はこれらの最大値を選ぶ問題であった。すなわち、 $L_i = \max(X_1, \dots, X_i)$ ,  $1 \leq i \leq n$  とすると

$$P\{X_{\tau^*} = L_n\} = \sup_{\tau} P\{X_{\tau} = L_n\}$$

となる停止時刻  $\tau^*$  を求める問題であった。それに対して、この論文は必ずしも最大値である必要はなく、最大値に十分近ければよいと考えられる。すなわち、 $L_n$  に関する許容関数  $\rho(L_n)$  を導入し、選んだ評価値と最大値  $L_n$  の差が許容範囲内であればよい。つまり

$$P\{X_{\sigma^*} \geq L_n - \rho(L_n)\} = \sup_{\sigma} P\{X_{\sigma} \geq L_n - \rho(L_n)\}$$

となる停止時刻  $\sigma^*$  を求める問題である。

- (b) **Tamaki**<sup>92)</sup>:  $m$  個のマイナスボール (ボールに “-1” の印) と  $p$  個のプラスボール (ボールに “+1” の印) が入っている壺から非復元抽出で毎時 1 個ずつボールを取り出し,  $X_j$  で時刻  $j$  に取り出されたボールの値を示す。すなわち,  $j$  番目のボールがマイナスボールかプラスボールかによって,  $X_j = -1$  あるいは  $X_j = 1$  とする。  $Z_0 = 0$ ,  $Z_n = \sum_{j=1}^n X_j$ ,  $1 \leq n \leq m+p$  で確率過程  $\{Z_n\}_{n=0}^{m+p}$  を定義し, この過程の最大値で停止する確率を最大にする問題である。

# 索引

<b>【い】</b>		<b>【さ】</b>		<b>【た】</b>	
一様分布	150	最後の成功	106	短縮型ルール	92
一様有界	242	最小期待順位	11	単調ケース	56
一般化された一様分布	151	最小値分布	153, 154	単調ルール	165, 176
<b>【か】</b>		最大成功確率	7	単峰	47
完全情報型	2	最適停止時刻	10	<b>【て】</b>	
完全情報型期間問題		最適停止問題	1	停止ルール	242
	195, 197, 199	最適方程式	6, 25	<b>【と】</b>	
完全情報型最良選択問題		最適ルール	7	動的計画法	6
	16, 198	最良選択問題	3	トレーニングサンプル	68
完全情報型順位最小化問題		坂口モデル	52	<b>【は】</b>	
	24	<b>【し】</b>		箱型領域	217
完全情報型モデル	16	しきい値	7	パレート分布	130
完全情報型問題	142, 163	しきい値ルール	7	範囲	180
ガンマ関数	116	事後分布	116	半無限平面	215
<b>【き】</b>		指示関数	25	<b>【ひ】</b>	
幾何分布	155	指数積分関数	18	秘書問題	1
期間問題	191	事前分布	116, 142	非斉次ポアソン過程	95
強度関数	95	順位最小化問題	3	<b>【ふ】</b>	
共役分布族	116	状態	54, 108, 143	部分情報型最良選択問題	180
拒否	32	勝利	107	部分情報型問題	190
<b>【く】</b>		<b>【せ】</b>		<b>【へ】</b>	
ゲーゴル	128	成功	17, 33, 68	平均値関数	96
グループ面接	111, 122	セカンドベスト	65, 216	平面ポアソン過程	213
<b>【こ】</b>		絶対順位	1	ベターな応募者	188
交換可能	129	漸近挙動	3, 72	<b>【ま】</b>	
交互対応	206	<b>【そ】</b>		マルチンゲール	16, 241
候補者	6, 17, 33	相対順位	1		
候補者選択問題	46	相対ベスト	17		
古典的秘書問題	1, 8				

マルチンゲール停止定理 16, 242  【む】 無記憶停止ルール 27 無限問題 73, 85 無情報型 2 無情報型期間問題 191 無情報型最良選択問題 6 無情報型順位最小化問題 10 無情報型モデル 3	無情報型問題 142, 143  【め】 メモリ 79 メモリサイズ 80 メモリサイズ 1 の問題 80 メモリサイズ 1 の NIRM 82  【ゆ】 有限問題 73	【り】 リコール 32 臨界曲線 218 臨界値 18  【れ】 レンジ 180  【わ】 割引を考慮した NIBC 48
--	--	--

【B】 BL 法 218 Bruss の連続時間モデル 160  【D】 DUR 197, 199  【E】 $e^{-1}$ -ルール 163 $e_F^{-1}$ -ルール 163  【F】 Ferguson の生成ルール 130 Ferguson の秘書問題 129 FI モデル 16 FIBC 16, 120 FIRM 24 FL 法 217  【G】 Gnedin の生成ルール 135  【K】 $k$ 期のオッズ 107	【M】 Moser 問題 27  【N】 NHPP 95 NI モデル 3 NIBC 6 NIRM 10  【O】 Odds 定理 107 OLA ルール 56  【P】 PET 180 Petruccielli モデル 32 Petruccielli–Porosiński–Samuels パラドックス 214 POR 142, 198 Porosiński モデル 164 PPP 213 Presman and Sonin モデル 143	【R】 Renyi の補題 6 Robbins 問題 27  【S】 Samuel–Cahn モデル 143, 157, 175 Smith の拒否モデル 32 Sum–the–multiplicative–odds 定理 123 Sum–the–odds 定理 107, 114  【T】 $t$ -ルール 163 Tamaki の拒否モデル 40  【X】 $x$ -ルール 160  【Y】 Yang のリコールモデル 32 ~~~~~ 【数字】 1–sla ルール 54, 55, 108, 140
---	--	--

—— 著者略歴 ——

1972年 名古屋工業大学工学部計測工学科卒業  
1974年 大阪大学大学院基礎工学研究科博士前期課程修了（数理系専攻）  
1977年 大阪大学大学院基礎工学研究科博士後期課程退学（数理系専攻）  
1977年 追手門学院大学講師  
1980年 工学博士（大阪大学）  
1980年 追手門学院大学助教授  
1986年 愛知大学助教授  
1988年 愛知大学教授  
2019年 愛知大学名誉教授

最良選択問題の諸相 —— 秘書問題とその周辺 ——

Various Aspects of the Best Choice Problem

—— The Secretary Problem and Its Related Areas —— © Mitsushi Tamaki 2023

2023年7月12日 初版第1刷発行

検印省略

著者 たま き みつ し  
置 司  
発行者 株式会社 コロナ社  
代表者 牛来真也  
印刷所 三美印刷株式会社  
製本所 有限会社 愛千製本所

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社  
CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844 ・ 電話 (03) 3941-3131(代)

ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-02840-9 C3355 Printed in Japan

(金)



JCCOPY < 出版者著作権管理機構 委託出版物 >

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構（電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jccopy.or.jp）の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。