

# 代替案の絞り込み過程を組み込んだ ブランド選択行動分析

勝又壮太郎, 阿部 誠

## 1. はじめに

企業のマーケティング努力によって、市場には多数のブランドが投入され、それぞれが異なる特性をもち、異なる便益を提供している。消費者はこれらのブランドの中から自分の選好に最も近い製品を選択することができる。市場にある入手可能なブランドが多ければ多いほど、消費者はより自分のニーズに合った製品を入手できる可能性が高くなり、消費者にもたらされる便益は高くなると考えられる。しかし、現実には、消費者は全てのブランドを評価して合理的な意思決定によって選択を行っているわけではない。市場に100種類以上の入手可能なブランドがあったとしても、その全てを多様な観点から評価し、順位付けすることは、一個人には困難である。経済学が仮定する合理的な選択行動を人間が実行することは、現実には不可能といえる。この「合理的」意思決定行動の非現実性については、早くは文献[1]によって指摘されている。文献[1]は多数の代替案や問題が存在する現実世界における、「合理性の限界」を指摘し、個人の選択行動についても、代替案の全てを検討するのではなく、選択できる集合の一部分について検討し、その中で最も好ましい代替案を選択するというモデルを提示している。

文献[1]が提示した「合理性の限界」は、現実的で優れた概念であり、その後の消費者行動研究にも影響を与えているが、この概念をそのまま消費者選択行動の実証分析に活用することは簡単ではない。例えば、ID付きPOSデータをはじめとする行動データから消費者行動の分析を行う場合、入手可能な情報は「購買

者」, 「購買日時」, 「購買ブランド」という観測された結果だけであり、その背後にある「検討したブランドの集合」の情報を得ることはできない。消費者が「絞り込まれた代替案の集合」を内部で構成していることはおそらく明らかであるが、この「限定された」集合は、個人の内部に存在するものであるため、観測することが非常に難しい。消費者に直接質問するなどの能動的な方法をとらなければ、企業側がその情報を入手することはできないのである。

このような、代替案の絞り込みという消費者の内的な過程を、観測される行動データから推測しなければならないという大きな問題が障害となり、消費者の内的状態を考慮した行動データの分析はいまだ発展途上である。この問題に取り組むいくつかの研究が提案されてはいるものの、解決されなければならない点が多い。このような背景を踏まえ、本研究では、行動データを用いた、代替案の絞り込み過程を組み込んだブランド選択モデルの構築を試みる。行動データから得られる情報のみを用いて、繰返し購買の中で蓄積されている消費者個人の知識や経験が構成する代替案の集合を仮定し、その集合を基軸とした選択モデルを構築するものである。

## 2. ブランド選択の構造

### 2.1 代替案の絞り込み

文献[1]が指摘したように、個人は全てのブランドを評価しているわけではない。この仮定をおいた行動モデルは、消費者行動研究の分野でも進展している。ブランドの絞り込みは、消費者の内的な処理によって行われているものであり、消費者の内面的な処理モデルの発展とともに進展してきた。初期の研究では、文献[2]がS-O-R型の購買意思決定モデルを提案しているが、その中で「想起集合 (evoked set)」という概念が提示されている。想起集合は、購買時点において消費者の内面に想起されるブランドの集合であり、

かつまた そうたろう  
長崎大学 経済学部  
〒850-8506 長崎市片淵4-2-1  
あべ まこと  
東京大学 大学院経済学研究科  
〒113-0033 文京区本郷7-3-1

市場に存在する全ブランドの部分集合になる。また、段階的なブランドの絞り込みを考慮したモデルも提案されている。例えば、文献[3]においては、全体集合>知名集合>考慮集合>選択集合>選択という多段階の構造を提案している。段階を進むにつれて、集合のサイズは小さくなるのである。また、文献[4]では、図1のような入手可能集合>知名集合>処理集合>想起集合>選択という構造を提案している（Brisoux と Laroche の概念図と呼ばれている）。

消費者行動研究で提案されたこのような集合の絞り込み過程は、概念的にも妥当性があり、消費者の内面的変化を精緻に切り分けることに成功している。しかしながら、これらのモデルも、中間段階が観測できないという大きな問題を持っている。入手可能集合は、店舗の品揃えともとらえることができ、これは企業が決定するものであるため、容易に観測することができる。また、選択したブランドも観測することができ、購買履歴として蓄積されている。しかしながら、中間の知名集合、処理集合、考慮集合は、直接消費者に聞かなければ観測することはできない。これらの絞り込まれた集合を、ID 付き POS データなどの行動データから得るためには、何らかのかたちで推定する必要がある。

いくつかの研究では、これを部分的に解決することに成功している。例えば、文献[5]では、既存のブランド選択モデルを拡張し、考慮集合の構成そのものを確率的に推定するモデルを提案している。現在でも多くの研究が文献[5]の考慮集合概念を導入してモデルを構築している。しかし、このモデルには実務上かなり大きな欠点がある。代替案の数が多くなると、全体集合  $\Omega$  のサイズ  $n(\Omega)$  に対して、考慮集合のパターンが  $2^{n(\Omega)} - 1$  に膨れあがってしまうのである。例えば、単純に 10 ブランドを分析対象とすると、考慮集合のパターンは 1,023 になる。実際に市場に流通している

ブランドは、製品カテゴリーによっては数十から数百にもなり、このような製品カテゴリーで、全ブランドを含めた分析をすることは非常に困難になる。

絞り込まれた集合の観測は難しく、確率的に推定する手法はあるが、これも活用は限定的である。そこで、次節では、消費者の情報探索行動に着目し、ブランドのスクリーニングを行っていく動的な過程から、絞り込まれた集合の形成を動的にとらえる方策を考えていく。

## 2.2 情報探索観点からの動的選択過程

消費者が購買に至るまでの内的過程において、製品・ブランドの情報を探索するという行動は、消費者情報処理モデルに組み込まれている。情報処理モデルは、Howard-Sheth モデル[2]をはじめとする S-O-R 型モデルをその理論的基礎とし、消費者の能動性をより強調したものである。情報処理モデルが表現しているのは、自身の内部に何らかの目標を持っており、その目標と現実のギャップ、すなわち課題が発生したとき、それを解決することのできる手段が何なのか、能動的に情報を収集するという消費者である。初期の代表的な研究としては、文献[6]によるモデル（Bettman モデル）があり、このモデルの枠組みはその後の研究でも踏襲され、例えば、文献[7]では CDP（Consumer Decision Process）モデルを提示している。

Bettman モデルにおいても CDP モデルにおいても共通しているのは、「消費者は問題解決の際に、まず自分の記憶を頼りに評価を行い、それが満足行く水準に達しなかったときには、外部の情報を探索する」、「購買後に製品を消費することで学習し、（長期）記憶へのフィードバックがかかる」という2つの構造を持つことである。内部探索・外部探索・学習について構造を整理すると、次のような過程となる。まず、消費者は問題解決の際に自らの記憶や経験からの「内部探索」を行い、必要に応じて外的情報を参照する「外部探索」を行う。そして購買後は記憶や経験を拡大する「学習」が行われ、次回、その経験や記憶から「内部探索」が行われる。すなわち、内部探索によって満足水準に達するブランドがなかったときには、引き続いて外部探索を行い、満足水準を超えるブランドを選択するという構図をとらえることができる。フィードバック機構に関しては、S-O-R モデルのひとつである Howard-Sheth モデルにおいても、出力としての「購買」から学習構成概念の「ブランド理解」へのフィードバックがモデルに組み込まれており、購買・利用経

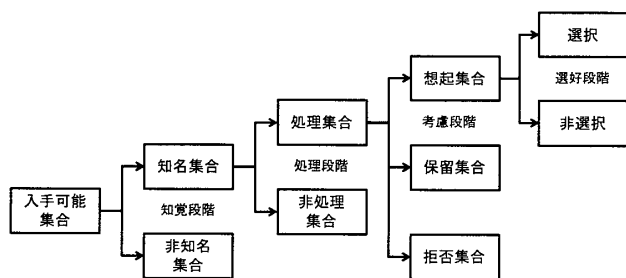


図1 Brisoux と Laroche の概念図

験による消費者の内的なブランド情報の更新を想定している。そこで、次節では、フィードバック機構による学習プロセスを、行動データから再構成する方策を考えてみたい。

### 2.3 反復的な購買による経験と学習

行動データのひとつであるID付きPOSデータとは、ブランドの購買履歴が、個人のIDを添えて時系列に記録されているものである。ただし、このデータからは、消費者の内的な情報処理過程を観察することはできず、購買の結果のみが、唯一、消費者の意図を反映したデータとして観測できる。すなわち、行動データから消費者の情報処理を考察するということは、探索行動と学習を観測できる結果から解釈するという他にない。

ある時系列行動データがあったとき、消費者*i*によって、任意の*t*時点でブランド*j*が購買されたとする。このブランド*j*が、*t*以前にも購買されていた記録があるとき、消費者はこのブランドに関する知識を持っていることは確かである。すなわち、かつて購買した経験のあるブランドは、内部探索によって購買されているといえる。逆に、消費者*i*がこれまで購買したことのないブランド*k*を購買したときには、これは外部探索によって購買されたと考えることが妥当である。そして、外部探索によって購買されたブランド*k*は、購買・使用を経て消費者に評価され、記憶に貯蔵されることになる。すなわち、次回以降の購買においては、ブランド*k*は内部探索によって検討される代替案になる。すなわち、*t*時点で、*t*より前に購買が観測されたブランドの集合は、記憶に貯蔵された代替案の集合に含まれているといえる。この、記憶に貯蔵された代替案の集合は、一定期間の行動データがあれば、そこから得ることが可能である。ただし、この集合は、厳密には「購買を検討している集合」ではなく、「考慮集合」、「選択集合」などとは質的な差異がある。また、この集合は学習・記憶によるフィードバック機構の後に形成される集合であるために、図1の概念図でいえば、「選択」の後段階に存在する集合といえる。図1の概念図は1回の購買現象を説明しているのみであるために明示的には示されていないが、この集合形成の過程は、何回も購買を行う場合、その度に繰り返されるものである。すなわち、この集合は反復的な購買行動を想定したときに初めて意識されるものであり、これまでの購買過程モデルにおける代替案の集合とは概念上別の位置にある集合であるため、この集合を新

たに「経験集合」と呼ぶこととする。経験集合は「反復的な購買行動の中で形成され、消費者の情報処理過程において内部探索を行う対象となるブランドの集合のうち、購買・利用経験のあるブランドの集合」と定義する。

経験集合という概念を導入することによって、観測できる情報から、一連の消費者情報処理モデルのフィードバック機構を明示的に組み込み、反復的な購買プロセスを動的に解釈することができるようになる。図2は、ここで再構成した、反復購買における探索と学習のモデルを示している。

## 3. 経験集合の購買モデル

### 3.1 経験集合の導入

消費者の長期記憶には、特定のブランドに対する評価が貯蔵されている。これは、過去の購買経験から構成されるものであり、前節で「経験集合」と定義した。本節では、この経験集合を用いた購買構造を定式化する。

まず、個人*i*が*t*回目の購買機会において保持している経験集合を $E_{it}$ とおく。経験集合は時間の経過とともに変化していくものであるため、*t*が添えられる。また、経験集合のサイズ、すなわち経験集合に含まれているブランドの数は $n(E_{it})$ とおく。ここでは、経験集合の変化と、それに伴う観測変数についての関係を図2に沿って定式化する。

個人*i*が時間*t*に購買を意図したとき、個人は購買前に内部探索を行う。このとき、内部探索の対象となるのは*t*-1回目の購買後に学習（ブランド理解）を

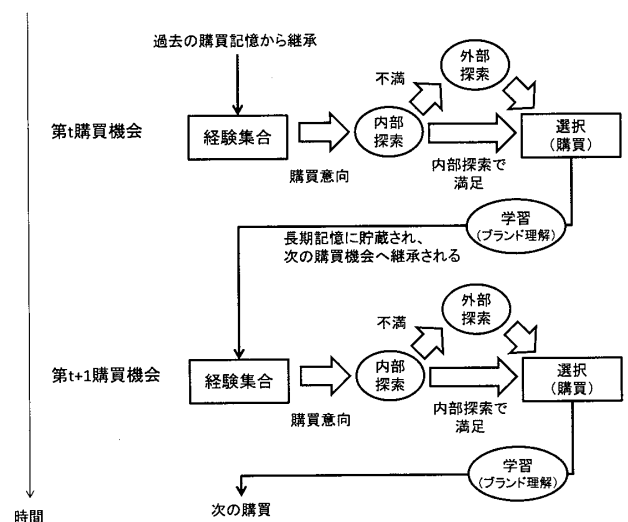


図2 反復的な探索と学習のモデル

経て形成され、継承された経験集合  $E_{it}$  である。ここで、 $t$  回目の購買において  $E_{it}$  に含まれているブランドを購買した場合、個人  $i$  は内部探索によって満足できるブランドを見つけ、それを購買したととらえることができる。このとき、時間  $t+1$  における経験集合は  $t$  における経験集合と等しく、 $E_{i,t+1}=E_{it}$  である。次に、個人  $i$  が  $t$  回目の購買機会において内部探索で満足いく水準の代替案を見つけることができなかつたとき、この個人は外部探索によって新しい代替案を見つけ出す。経験集合  $E_{i,t}$  に含まれていないブランドが購買されたときに、このような情報処理が消費者の内部で行われたととらえることができる。ここで、経験集合  $E_{it}$  に含まれていないブランド  $j$  が時間  $t$  に購買されたとき、経験集合  $E_{i,t+1}$  は  $E_{it}$  に含まれているブランドに加えて  $j$  を含んだ集合となる。外部探索の結果、経験集合が拡大したのである。

ここで、観測される変数を定義する。まず、個人  $i$  によって  $t$  回目の購買機会にブランド  $j$  が購買されたとき、購買の変数  $y_{itj}=1$  とおく。ブランド  $j$  は経験集合に含まれている ( $j \in E_{it}$ )。また、経験集合に含まれていながらも購買されなかつたブランド  $k$  ( $k \neq j, k \in E_{it}$ ) について、 $y_{itk}=0$  である。さらに、本研究では、外部探索についても変数を定義する。外部探索は、購買されたブランド  $j$  が前期の経験集合に含まれていなかったときに観測される。すなわち、 $j \notin E_{it}$  のときに、外部探索が発動しているため、このときの外部探索変数  $z_{it}=1$  とおく。このとき、ブランド  $j$  について初回の購買が観測されることになるが、ここでは経験集合に入っていないために、ブランド  $j$  の購買確率の推定には用いない。経験集合の定義で述べたように、経験集合に含まれているブランドは、内部探索を行うブランドのうち、特に、購買・利用経験があるものである。すなわち、そのブランドを初めて選択した購買時点ではなく、購買後に経験集合は拡大する。

こうした内部探索、外部探索、経験集合の構造を示したのが、表1である。初期状態（時間0）には、ブランド  $a$  のみが経験集合に入っている個人が、外部探索を繰り返すことで経験集合を拡大させていく過程を表現している。ここで、初回の購買は(1)とされている<sup>1</sup>。

### 3.2 モデルの定式化

ここでは、観測される  $y$  と  $z$  を説明する構造を仮

表1 経験集合と外部探索の構図

購買機会 $t$	ブランド購買 $a$ $b$ $c$	経験集合 $E_{it}$	外部探索 $z_{it}$
0		{ $a$ }	
1	1	{ $a$ }	0
2	1 (1)	{ $a$ }	1
3	0 0	{ $a, b$ }	0
4	0 0	{ $a, b$ }	0
5	0 0 (1)	{ $a, b$ }	1
6	1 0 0	{ $a, b, c$ }	0
7	0 1 0	{ $a, b, c$ }	0
8	1 0 1	{ $a, b, c$ }	0

定する。得られた観測値は離散的な値をとるために、これを説明するためには離散選択モデルを利用する必要がある。そこで、本研究では、連続的な潜在変数を仮定したモデルを構築する。これは、文献[8]によって提案された方法で、マルコフ連鎖モンテカルロ (MCMC) 法によってパラメータの推定を行う点の特徴である。MCMC 法には後述するいくつかの利点があり、本研究でも、MCMC 法によってパラメータの推定を行う。まず、ブランド購買の観測値  $y_{itj}$  について、連続的な潜在変数  $y_{itj}^*$  を導入する。 $y_{itj}$  と  $y_{itj}^*$  は、次のような関係にある。

$$y_{itj} = \begin{cases} 1, & \text{if } y_{itj}^* > 0, \\ 0, & \text{if } y_{itj}^* \leq 0. \end{cases} \quad (1)$$

潜在変数  $y_{itj}^*$  は、ブランド  $j$  に対して個人  $i$  が時間  $t$  に評価している効用ととらえることができる。ブランド  $j$  が購買されるためには、潜在変数  $y_{itj}^*$  が一定の水準を超えている必要があり、その閾値を0とおいている。ここで、購買機会  $t$  において、経験集合に入っている全てのブランドの潜在変数の値が0を下回れば、この購買機会には何も購買されず、2個以上のブランドの潜在変数の値が0よりも大きければ、そのブランドは全て購買されるというモデルである。

ブランド購買の観測値  $y_{itj}$  と同様に、外部探索の観測値  $z_{it}$  についても、連続的な潜在変数  $z_{it}^*$  を導入する。 $z_{it}$  と  $z_{it}^*$  の関係は以下の通りである。

$$z_{it} = \begin{cases} 1, & \text{if } z_{it}^* > 0, \\ 0, & \text{if } z_{it}^* \leq 0. \end{cases} \quad (2)$$

続いて、こうして定義したブランド選択行動  $y_{itj}$  と外部探索行動  $z_{it}$  の潜在変数を説明する構造を考える。

<sup>1</sup> この例では、購買機会があっても対象の製品カテゴリーからブランドをひとつも購買しないというケースを含めて分析対象としており、また、1回の購買機会に2つ以上のブランドを購買するケースも含んでいる。

まず、時間  $t$  に個人  $i$  が接触するセールスプロモーション等に対する変数を  $x_{it}$  とおき、反応度を  $\beta_{ij}$  とおく。ここから、ブランド選択を説明するモデルは、次のように表すことができる。

$$y_{it}^* = x_{it}'\beta_{ij} + \varepsilon_{it}, \varepsilon_{it} \sim \mathcal{N}(0, 1), j \in \varepsilon_{it}, t=1, \dots, T_i \quad (3)$$

次に、外部探索については、外部探索傾向を説明する変数  $w_{it}$  とパラメータ  $\gamma_{it}$  から、次のように構成する。

$$z_{it}^* = w_{it}'\gamma_i + \zeta_{it}, \zeta_{it} \sim \mathcal{N}(0, 1), t=1, \dots, T_i \quad (4)$$

ここで、各パラメータに上部の構造を設定し、階層化によってパラメータ間の関係を検討する。階層化したモデルの推定は、MCMC 法ならば可能であり、これも本研究で MCMC 法を用いる理由のひとつである。個人はパラメータ  $\beta_{ij}, \gamma_i$  をもち、これらのパラメータの値は、個人のデモグラフィック属性によって一定程度説明可能であると考えられる。そこで、個人ごとに推定されているパラメータを個人属性によって説明する。個人属性の変数を  $v_i$  とおき、それにかかるパラメータを  $\Delta_j$  および  $\Lambda$  とおいて、次のような構造を仮定する。

$$\beta_{ij} = \Delta_j v_i + \xi_{ij}, \xi_{ij} \sim \mathcal{N}(0, \Psi), \quad (5)$$

$$\gamma_i = \Lambda v_i + \xi_i, \xi_i \sim \mathcal{N}(0, \Phi). \quad (6)$$

具体的な説明変数は、次節でデータの概要を説明したあとに解説する。また、事前分布・事後分布・推定法などは付録に記載している。

## 4. 実証分析

### 4.1 データ概説

前節で解説したモデルを、実際のデータを用いて検証する。本研究で用いるデータは、経営科学系研究部会連合協議会が主催する平成 22 年度データ解析コンペティションにおいてカスタマー・コミュニケーションズ株式会社から提供された、ドラッグストアの購買履歴データである。期間は 2008 年 1 月 1 日から 2009 年 12 月 31 日までの 2 年間である。対象とする製品カテゴリーはシャンプーである。シャンプーは低価格の必需品であるため、反復的に購買される製品であり、本研究のモデルを実証するためには好適であるといえる。

### 4.2 実証分析の設定

#### 4.2.1 分析期間・分析対象ブランド

まず、分析対象の期間について説明する。前節で提

表 2 分析対象のブランド

順位	ブランド名	順位	ブランド名
1	ラックス	11	h & s
2	パンテーン	12	アジェンス
3	TSUBAKI	13	いち髪
4	メリット	14	シーブリーズ
5	エッセンシャル	15	ヴィダルサスーン
6	ダヴ	16	海のうるおい薬
7	スーパーマイルド	17	水分へアパック
8	ソフトインワン	18	トニックシャンプー
9	PB(プライベートブランド)	19	植物物語
10	モッズヘア	20	オクト

案したモデルを推定するためには、分析対象の期間の前に一定の時間をおき、経験集合の初期状態を設定する必要がある。そこで、本研究では、2 年 (24 カ月) 間のデータのうち、はじめの 9 カ月間を経験集合の形成期間としておき、その後を分析対象の期間とする。経験集合の初期状態には、はじめの 9 カ月間に購買が観測されたブランドが含まれている。また、モデルによる予測精度の検証のために、各個人の最後の購買機会のデータを取りおき、これを検証用データとする。

次に、分析対象とするブランドは、全期間を通して購買数が多かった上位 20 ブランドとした。モデルの構造として、対象とするブランド数が多くなっても計算負荷が大きく増加することはないが、あまり購買数が少ないブランドだと、推定結果の信頼性が低下するため、20 ブランドを対象とした。具体的なブランド名は、表 2 に記載している。

#### 4.2.2 購買機会と分析対象顧客

まず、購買機会を定義する。前節までのモデル構築で、本研究で扱うブランド選択モデルは、同じ購買機会に 2 個以上のブランドが購買される場合を想定し、また、購買機会に 1 つもブランドが購買されていない場合も想定している、多変量選択 (Multivariate Choice) 構造としている。これを踏まえて、購買機会を定義する。購買機会は、来店・非来店を問わないカレンダー通りの日にちとすることも、来店日とすることも可能だが、本研究では、「シャンプーカテゴリーの購買があった日 (カテゴリー購買が観測された日)」を購買機会ととらえる。分析対象の顧客を抽出する条件は、経験集合の形成期間・分析対象期間ともに、対象のブランドを 3 回以上購買していることを条件とした。この条件に合致した顧客から無作為に 400 人を選び出し、分析対象とする。

### 4.2.3 説明変数

続いて、説明変数について解説する。まず、ブランド選択の説明変数  $x$  に含まれる変数について説明する。対象の店舗では、毎月1日および20日に全製品を値引きするセール日がある。そこで、これをダミー変数の形で説明変数の1つとして用いる。また、同じく説明変数として、土日・祝祭日をダミー変数の形で用いる。次に、外部探索を説明する変数  $w$  について説明する。 $w$  には、 $x$  と同様に、セール日ダミーと土日・祝祭日ダミーを組み込むため、本研究では  $x$  と等しくなる。

## 5. 推定結果の考察

### 5.1 モデルの収束と予測

まずは、推定したモデルの収束を判定する。収束の判定においては、文献[9]で提案された判定法を用いている。前半10%と後半50%のサンプルについて検定を行ったところ、全てのパラメータは収束していることが確認できた。

次に、予測のためにとりおいたデータを用いて、モデルの予測能力を検討する。ここでは、 $t = T_i + 1$  回目の購買確率を計算して、予測用のスコアとする。まず、顧客  $i$  の時間  $t$  におけるブランド  $j$  の購買確率の推定値  $\hat{p}_{itj}$  は、ブランド  $j$  が経験集合  $\varepsilon_{it}$  に入っているとき、次の式から求めることができる。ここで、 $\hat{\beta}_{ij}$  は  $\beta_{ij}$  の事後平均値であり、 $\Phi(\cdot)$  は標準正規分布の分布関数である。

$$\hat{p}_{itj} = \Pr(\hat{y}_{itj} = 1) = \Phi(x'_{it}\hat{\beta}_{ij}), \text{ if } j \in \varepsilon_{it} \quad (7)$$

同様に、外部探索を行う確率  $q_{it}$  は、次の式から計算することができる。

$$\hat{q}_{it} = \Pr(\hat{z}_{it} = 1) = \Phi(w'_{it}\bar{v}_i) \quad (8)$$

経験集合に入っていないブランドについては、予測購買確率をそのまま求めることができない。そこで、事前分布から  $\beta_{ij}, j \in E_{i,T_i}$  を推定し、それを用いて購買確率を計算する。 $\beta_{ij}, j \notin E_{i,T_i}$  の推定値は、 $\Delta_j$  の事後平均  $\bar{\Delta}_j$  および個人属性の変数  $v_i$  を使って求めることができる。 $\beta_{ij}, j \in E_{i,T_i}$  の推定値を  $\hat{\beta}_{ij} = \bar{\Delta}_j v_i$  とおいて、このパラメータと  $x_{it}$  から、経験集合外のブランドの選択確率を推定する。ただし、この購買が実際に観測されるためには外部探索が発動することが必要なので、外部探索確率  $\hat{q}_{it}$  を掛ける。

$$\hat{p}_{itj} = \hat{q}_{it} \Phi(x'_{it}\hat{\beta}_{ij}), \text{ if } j \notin E_{it} \quad (9)$$

表3 予測結果

ブランド名	購買数	ROCスコア	ブランド名	購買数	ROCスコア
ラックス	62	0.935	h & s	24	0.978
パンテーン	43	0.855	アジエンズ	13	0.932
TSUBAKI	34	0.936	いち髪	8	0.976
メリット	38	0.885	シーブリーズ	6	0.993
エッセンシャル	25	0.920	ヴィダルサスーン	15	0.927
ダヴ	25	0.933	海のうるおい薬	13	0.931
スーパーマイルド	15	0.821	水分ヘアパック	5	0.996
ソフトインワン	18	0.885	トニックシャンプー	10	0.992
PB	17	0.824	植物物語	8	0.999
モッズヘア	11	0.987	オクト	3	0.995
			外部探索	43	0.806

予測能力の判定においては、ROC (Receiver Operating Characteristics) 曲線とROCスコアを用いる。ROCスコアは、0から1の間の値をとり、その値が1に近いほど予測能力が高いといえる。また、完全にランダムに予測を行った場合には統計的に0.5をとるため、0.5を上回っていれば予測能力があることを示している。詳細は文献[10]を参照のこと。得られた結果を表3に示している。

表にある「購買数」は、予測用のサンプルで購買が観測された数である。表をみると分かるように、ROCスコアはともに非常に高いことが分かる。各ブランドを選択する傾向が高い個人を高い精度で選別することができているといえる。特に、購買数が少ないブランドについても、高い精度で購買者を判別し、予測できている。また、外部探索の予測において得られた値も高く、モデルで外部探索確率が高いと判定された個人が実際に高い頻度で外部探索を行っているということがいえる。

このROCスコアの算出のために用いるROC曲線について、代表的なものを図3に挙げている。ROC曲線は、中央の点線を超えれば予測能力があるとされ、図の左上に寄っているほど予測精度が高いことを示している。これを見ても分かるように、非常に高い精度で予測ができており、実用的なモデルであるといえる。

次節以降では、得られた個々のパラメータを考察し、マーケティング施策への活用を考える。

### 5.2 パラメータの考察

本節では、得られたパラメータの考察を行うが、まず、パラメータの意味を考えたい。本モデルで得られたパラメータ  $\beta_{ij}$  は、経験集合に入っているブランドの購買傾向を表すものである。そのブランドの利用経験がある消費者の集団から推定したパラメータであり、切片項に掛るパラメータである  $\beta_{ij1}$  は、消費者  $i$  の「再購買意向」を示していると解釈することができる。この値が高いブランドは、一度購買した消費者が再び

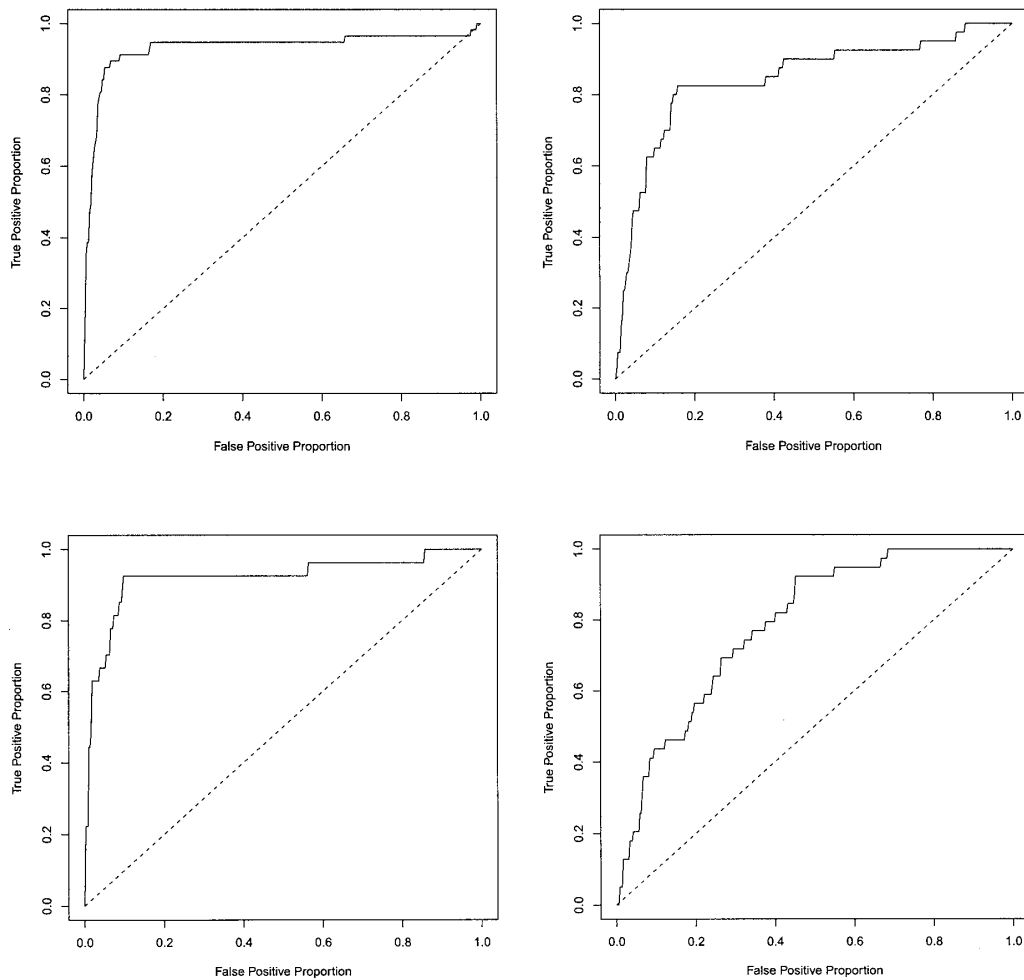


図3 ROC 曲線 (左上) ラックス, (右上) パンテーン, (左下) TSUBAKI, (右下) 外部探索

表4 ブランド切片  $\bar{\beta}_{ij1}$  の平均と経験集合内の人数

ブランド名	平均値	標準偏差	人数	ブランド名	平均値	標準偏差	人数
ラックス	-0.53	0.49	150	h & s	-0.97	1.84	62
パンテーン	-1.25	0.62	155	アジエンス	-1.72	0.74	58
TSUBAKI	-1.88	2.42	107	いち髪	-3.25	6.47	44
メリット	-1.66	0.49	107	シーブリーズ	-2.09	0.65	53
エッセンシャル	-1.15	0.71	81	ヴィダルサスーン	-1.79	0.56	63
ダヴ	-0.95	0.62	71	海のうるおい薬	-2.13	0.60	51
スーパーマイルド	-2.35	0.65	84	水分ヘアパック	-2.74	1.16	37
ソフトインワン	-0.62	0.71	43	トニックシャンプー	-1.49	0.76	31
PB	-1.75	0.71	58	植物物語	-2.14	0.35	39
モッズヘア	-1.90	0.56	72	オクト	0.84	3.16	10

購買する傾向が高く、この値が低いブランドは、一度購買した消費者が再購買する傾向は低いと考えられる。

表4は、個人ごとに推定される各ブランドの切片の事後平均  $\bar{\beta}_{ij1}$  を  $i$  について集計した平均値と標準偏差を求めたものである。また、人数は、 $T_i$  時点でそのブランドが経験集合に入っていた消費者の数を示している。ブランド切片は経験集合に入っている消費者に対して推定されるので、これは平均および標準偏差を計算するときの標本数でもある。この表を見ると、ブ

ランド間で平均値と標準偏差にかなりの差異があることが分かる。例えば、「ダヴ」と「スーパーマイルド」は、経験集合に入っている消費者の数は同程度であるが、切片の平均値に差があることが分かる。また、「オクト」は、経験集合に入っている消費者はわずか10人だが、ブランド切片の平均は非常に高いことが分かる。広告や店頭プロモーションによって、1回の購買を促進させることももちろん重要であるが、長期にわたって利益やシェアを獲得していくためには、購買経験者に評価してもらい、再び購買してもらうことが非常に重要である。ブランド切片  $\beta_{ij1}$  は、ブランドが長期的に消費者の支持を集めることができるのかどうかを推測するための重要な情報であるといえる。

また、 $\Delta_j$  から、そのブランドがどのような属性の消費者から支持を集めているのかを確認することができる。表5は、ブランド「ラックス」についてのパラメータ  $\Delta$  の要約である。数値は事後平均で、\*はHPD基準で95%有意を示している<sup>2</sup>。「ラックス」は、

表5 ブランド「ラックス」の $\Delta$

	切片	性別	年齢
切片	-0.74	0.28	-0.01
セール日	-0.60	1.22*	-0.18
土日祝日	-1.07	-1.30*	0.68

セール日には女性，土日休日には男性に購買される確率が高くなるようだ。予測の節でも説明したように，このパラメータは，あるブランドについて，その購買経験がない，すなわち経験集合に入っていない消費者の「そのブランドが経験集合に入ったとき」の推定値を計算するために用いることができる。この推定値は，利用経験がある集団から得られた値なので，ブランドのレコメンデーションにも利用することができる。

## 6. 結論

本研究では，多数のブランドが存在し繰返し購買が観測される消費財市場における消費者のブランド選択行動を再解釈し，消費者個人の内面にある絞り込まれたブランドの集合を考慮した選択モデルを構築し，実証を行った。本研究の貢献として，以下の3点を挙げることができる。

第1に，ID付きPOSデータをはじめとした観察（行動）データを分析するための理論的な枠組みを構築したことである。消費者行動研究分野で提示されていた内部探索・外部探索・学習の概念を整理し，繰返し購買による動的な観察された購買行動を，消費者の情報処理の観点から再解釈した。これらの概念を整理し，モデルに取り入れたことで，より理論的に妥当なモデルを構築することができた。また，これまでの研究で概念として提示されていた考慮集合や処理集合といった，消費者の内部に存在する集合は直接質問するなどの能動的な手段でしか把握することができないが，経験集合という観察できる集合の概念を提示することで，実用的なモデルを構築することができた。本研究で提示したモデルは，これまでのブランド選択モデルと比較して，代替案の増加に対して頑健であり，市場に数十のブランドが存在しても，計算負荷を大きく増やさず分析することができる。

第2は，モデルの予測能力の高さである。本研究で提案したモデルは，予測用データを用いた検証から，

<sup>2</sup> HPDの計算は，文献[11]による方法を用いた。

非常に高い予測能力を持つことが示された。モデルは，経験集合に入っているブランドの選択確率が高くなるように設計されているが，このモデルで予測能力が高いということは，多くの消費者が経験集合に入っているブランドを優先的に選択する確率が高いことを示している。消費者がある程度検討するブランドの集合を絞り込んでいることは多くの先行研究で主張されているが，予測の結果から，過去に自分が購買したブランドを検討している可能性が高いということが示された。また，ブランド選択の予測だけでなく，外部探索の予測についても良好な結果を得ており，外部探索をしやすい消費者を高い精度で判別することに成功している。外部探索をしやすい消費者は，バラエティ・シーキング傾向が高い消費者ともいえる。このモデルを用いることで，消費者ごとの対象製品カテゴリにおける態度の差異を把握することができる。

第3は，モデルの拡張可能性の高さである。本研究で提案したモデルは，多くの部分でこれまでのブランド選択モデルの枠組みを踏襲している。また，階層化し，MCMC法によって推定を行っており，これまでの研究で提案されたモデルを柔軟に組み込むことができる。例えば，カテゴリ横断的なモデルや時系列項の導入など，多様な拡張を考えることができる。

今後の課題としては，次の2点を挙げる。第1は，より幅広い製品カテゴリを対象とした実証である。本研究ではシャンプーを対象に実証分析を行い，高い予測能力があることが確認できたが，他の製品カテゴリにおいても同様の予測能力を発揮できるのかを検討しなければならない。予測の結果から，経験集合という概念の妥当性はある程度示すことができたと考えられるが，他のカテゴリにおいても経験集合内から選択が行われるのか，検証しなければならない。第2は，モデルの精緻化である。例えば，本研究では，経験集合内のブランドが検討する代替案から外れる機構を組み込んでいない。分析対象の期間が短いので，本研究においては大きな問題は起きていないが，より長期間の分析を行うのであれば，検討する必要があるだろう。その他，実証を重ねていく上で必要があれば，適宜改良していくことが望ましい。

**謝辞** 本稿の執筆にあたり，経営科学系研究部会連合協議会およびカスタマー・コミュニケーションズ株式会社よりデータを貸与いただきました。また，西本章宏氏（小樽商科大学）から有益なコメントをいただき



ました。この場をお借りして御礼申し上げます。

## A. モデルの推定方法

### A.1 事前分布

ここでは、事前分布について解説する。以下、 $N$  は分析対象とした消費者の数、 $J$  は分析対象としたブランドの数、 $T_i$  は分析対象期間、 $K$  は説明変数  $x_{it}$  および  $w_{it}$  の次元、 $L$  は説明変数  $v_i$  の次元である。

まず、 $\Delta_i$  については、 $\Delta_j = \{\delta_{j1}, \dots, \delta_{jK}\}'$  となる  $L$  次元のベクトル  $\delta_{jk}$  に分割し、各ベクトルについて多変量正規分布の事前分布を仮定する。すなわち、 $\delta_{jk} \sim \mathcal{N}(\delta_0, \Sigma_{\delta_0})$ ,  $j=1, \dots, J, k=1, \dots, K$  とする。同様に、 $\Lambda$  についても、ベクトルに分割した  $\lambda_k$  について、事前分布を  $\lambda_k \sim \mathcal{N}(\lambda_0, \Sigma_{\lambda_0})$ ,  $k=1, \dots, K$  とおく。本研究では、 $\delta_0 = \lambda_0 = \mathbf{0}$ ,  $\Sigma_{\delta_0} = \Sigma_{\lambda_0} = 10^2 I_L$  とした。

つぎに、分散項  $\Phi$  および  $\Psi$  については、本研究では  $\Phi = \Psi$  としたので、以後  $\Psi$  について説明する。 $\Psi$  は対角行列とし、 $\Psi = \text{diag}(\psi_1, \dots, \psi_K)$  となる  $\psi_k$  について、事前分布を  $\psi_k \sim \mathcal{G}(n_0/2, n_0 S_0/2)$ ,  $k=1, \dots, K$  とおく。本研究では、安定した推定値を得るために、事前パラメータは  $n_0 = 10^4$ ,  $S_0 = 10^{-1}$  とした。

$\Psi$  を対角行列としたので、 $\beta_{ij}$  および  $\gamma_i$  の式は、スカラーの方程式に分解でき、 $k$  要素について、以下のように表記することができる。ここで、 $\mathcal{M}_j$  は、 $T_i$  時点でブランド  $j$  を経験集合に含んでいる消費者の集合である。

$$\beta_{ijk} = v_i' \delta_{jk} + \xi_{ijk}, \quad \xi_{ijk} \sim \mathcal{N}(0, \psi_k),$$

$$j=1, \dots, J, i \in \mathcal{M}_j, k=1, \dots, K, \quad (10)$$

$$\gamma_{ik} = v_i' \lambda_k + \eta_{ik}, \quad \eta_{ik} \sim \mathcal{N}(0, \psi_k), \quad k=1, \dots, K \quad (11)$$

### A.2 事後分布

$y_{itj}$  の事後分布：潜在変数  $y_{itj}^*$  は、文献[8]の方法に従い、切断正規分布から推定する。 $y_{itj}=1$  のとき、 $y_{itj}^* | \cdot \sim \mathcal{TN}_{(0, \infty)}(x_{it}' \beta_{ij}, 1)$  となり、 $y_{itj}=0$  のとき、 $y_{itj}^* | \cdot \sim \mathcal{TN}_{(-\infty, 0]}(x_{it}' \beta_{ij}, 1)$  となる。ここで、 $\mathcal{TN}_{(a, b)}(m, s)$  は区間  $(a, b)$  をとる、平均  $m$ 、標準偏差  $s$  の切断正規分布を示す。

$z_{it}$  の事後分布：これも  $y_{itj}^*$  と同様に、切断正規分布からサンプリングを行う。 $z_{it}=1$  のとき、 $z_{it} | \cdot \sim \mathcal{TN}_{(0, \infty)}(w_i' \gamma, 1)$  であり、 $z_{it}=0$  のときは、 $z_{it} | \cdot \sim \mathcal{TN}_{(-\infty, 0]}(w_i' \gamma, 1)$  となる。

$\beta_{ij}$  の事後分布：事後分布は、 $\beta_{ij} | \cdot \sim \mathcal{N}(b_1, S_{\beta 1})$ ,  $j=1, \dots, J, i \in \mathcal{M}_j$  になる。ここで、

$$S_{\beta 1} = (X_i' \bar{I}_{ij}^{(T)} X_i + \Psi^{-1})^{-1}, \quad b_1 = S_{\beta 1} (X_i' \bar{I}_{ij}^{(T)} Y_{ij}^* +$$

$\Psi^{-1} \Delta_j v_i)$  である。ただし、 $X_i = (x_{i1}, \dots, x_{iT_i})'$ ,  $Y_{ij}^* = (y_{i1j}^*, \dots, y_{iT_i j}^*)'$  であり、 $\bar{I}_{ij}^{(T)}$  は  $T_i \times T_i$  の対角行列で、その対角項  $(t, t)$  要素について  $j \in E_{it}$  ならば 1 をとり  $j \notin E_{it}$  ならば 0 をとる行列である。

$\gamma_i$  の事後分布：事後分布は、 $\gamma_i | \cdot \sim \mathcal{N}(g_1, S_{\gamma 1})$  になる。ここで、 $S_{\gamma 1} = (W_i' W_i + \Psi^{-1})^{-1}$ ,  $g_1 = S_{\gamma 1} (W_i' Z_i^* + \Psi^{-1} \Lambda_j v_i)$  である。ただし、 $W_i = (w_{i1}, \dots, w_{iT_i})'$ ,  $Z_{ij}^* = (z_{i1j}^*, \dots, z_{iT_i j}^*)'$  である。

$\delta_{jk}$  の事後分布：事後分布は、 $\delta_{jk} | \cdot \sim \mathcal{N}(d_1, S_{\delta 1})$  になる。ここで、 $S_{\delta 1} = (\phi_k^{-1} V' \bar{I}_j^{(N)} V + \Sigma_{\delta 0}^{-1})^{-1}$ ,  $d_1 = S_{\delta 1} (\phi_k^{-1} V' \bar{I}_j^{(N)} B_{jk} + \Sigma_{\delta 0}^{-1} \delta_0)$  である。ただし、 $V = (v_1, \dots, v_N)'$ ,  $B_{jk} = (\beta_{ijk}, \dots, \beta_{Njk})'$  であり、 $\bar{I}_j^{(N)}$  は  $N \times N$  の対角行列で、その対角項  $(i, i)$  要素について  $i \in \mathcal{M}_j$  ならば 1 をとり  $i \notin \mathcal{M}_j$  ならば 0 をとる行列である。

$\lambda_k$  の事後分布 事後分布は、 $\lambda_k | \cdot \sim \mathcal{N}(l_1, S_{\lambda 1})$  になる。ここで、 $S_{\lambda 1} = (\phi_k^{-1} V' V + \Sigma_{\lambda 0}^{-1})^{-1}$ ,  $l_1 = S_{\lambda 1} (\phi_k^{-1} V G_k + \Sigma_{\lambda 0}^{-1} \lambda_0)$  である。ただし、 $G_k = (\gamma_{1k}, \dots, \gamma_{Nk})'$  である。

$\psi_k$  の事後分布 事後分布は、 $\psi_k | \cdot \sim \mathcal{G}(n_1/2, n_1 S_1/2)$  になる。ここで、 $n_1 = n_0 + \sum_{j=1}^J \#(\mathcal{M}_j) + N$ ,  $n_1 S_1 = n_0 S_0 + \sum_{j=1}^J \sum_{i \in \mathcal{M}_j} (\beta_{ijk} - v_i' \delta_{jk})^2 + \sum_{i=1}^N (\gamma_{ik} - v_i' \lambda_k)^2$  である。ただし、 $\#(\mathcal{M}_j)$  は集合  $\mathcal{M}_j$  の要素数。

### A.3 初期値とシミュレーション設定

初期値について、潜在変数  $y_{itj}^*$  および  $z_{it}$  はすべて 0、パラメータ  $\beta_{ij}$ ,  $\gamma_i$ ,  $\delta_{jk}$ ,  $\lambda_k$  もすべて零ベクトルとしている。また、 $\psi_k$  は 1 としている。シミュレーションにおいては、全 15,000 回のイタレーションを行い、はじめの 5,000 回を除いた 10,000 回をサンプルとして採取している。

### 参考文献

- [1] Simon, H.A.: *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, Macmillan, New York, 1947.
- [2] Howard, J.A. and Sheth, J.N.: *The Theory of Buyer Behavior*, John Wiley & Sons, 1969.
- [3] Lapersonne, E., Laurent, G. and Le Goff, J.-J.: "Consideration Sets of Size One: An Empirical Investigation of Automobile Purchases," *International Journal of Research in Marketing*, 12, 55-66, 1995.
- [4] Brisoux, J.E. and Cheron, E.J.: "Brand Categorization and Product Involvement," *Advances in Consumer Research*, 17, 101-109, 1990.

- [5] Andrews, R. and Srinivasan, T.C.: "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32, 30-41, 1995.
- [6] Bettman, J.R.: *An Information Processing Theory of Consumer Choice*, Addison-Wesley Publishing Company, 1979.
- [7] Blackwell, R.D., Miniard, P.W. and Engel, J.F.: *Consumer Behavior 10th ed.* OH: Thomson Higher Education, 2006.
- [8] Albert, J.H. and Chib, S.: "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669-679, 1993.
- [9] Geweke, J.: "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," In J.O. Berger, J.M. Bernardo, A.P. Dawid and A.F.M. Smith (Eds.), *Bayesian Statistics 4 (pp. 169-194)*, Oxford University Press, 1992.
- [10] Blattberg, R.C., Kim, B. and Neslin, S.A.: *Database Marketing: Analyzing and Managing Customers*, Springer, 2008.
- [11] Chen, M.H., Shao, Q.M. and Ibrahim, J.G.: *Monte Carlo Methods in Bayesian Computation*, Springer, 2000.