

カテゴリーベース処理の選択行動モデルと 市場戦略への示唆

勝又 壮太郎, ウィラワン・ドニ・ダハナ, 中島 望

1. はじめに

現代の消費財市場は非常に激しい競争状態にある。とくに FPG (Frequently Purchased Goods) といわれる低価格消費財市場は新陳代謝が激しく、「狂騒的競争市場」とも評されている [1, 2]。競争の激しい FPG 市場で消費者の選択を勝ち得ていくためには、数多くの競合するブランドの中で、自ブランドと競合しているブランドを適切に把握しなければならない。消費者の選択を巡って競合しているブランドがわからなければ、競争優位性を獲得し、差別化を図っていくための適切な戦略を実行することができない。

FPG 市場をはじめとした市場における選択行動の分析には、ブランド選択モデルが用いられることが多い。とくに ID 付き POS データの分析においては、個人異質性を考慮したブランド選択モデルによって、消費者個人のブランド選好を高い精度で推定することに成功している [3]。しかしながら、高い予測能力をもつ個人異質性を考慮したモデルであっても、過度に競争的・流動的な市場の分析は困難であることが多い。第一の理由は、モデルが検討対象とするブランドの範囲に関する点である。通常のブランド選択モデルは、選択対象となるブランド数が固定され、対象ブランドは分析者が与えなければならない。現実の市場では時間の経過とともにブランドの増減があり、新ブランドや撤退するブランドがみられるが、選択対象の動的な変

化を仮定することは困難である。同様の理由から、新ブランドが多く発売される市場であっても、新ブランドの選択行動をモデルに組み込むことが難しい。第二の理由は、実際の消費者行動とモデルの仮定に関する点である。ブランド選択モデルは、各ブランドが選択される確率の背景に、個別ブランドに対する潜在的な効用を仮定しており、効用の相対的な高さに従って選択確率が決定されていくという構造をもつ。また、効用は製品の価格をはじめとするマーケティング変数によって説明され、一般には多属性態度モデルが仮定されている。すなわち、消費者が選択対象となるすべてのブランドについて検討し、効用をもっているという仮定になる。しかしながら、FPG 市場において、市場に投入されているブランドは数十から数百に上り、すべてのブランドを消費者が検討しているとは考えにくい。消費者は限定合理的な意思決定を行っていると考えられ、通常のブランド選択モデルをそのまま活用することは適切ではない。

上記 2 点の問題については、[4] によって一つの解決法が提案され、実際に数十のブランドを対象とした分析においても適切な結果を得ることが可能であることが示されている。しかし、消費者行動の観点からは、さらに第三の問題が存在している。消費者が個別のブランドではなく、より大括りのブランドの集合である「カテゴリー」単位で意思決定を行っている可能性である。FPG 市場において、消費者は、個別のブランドを検討する十分な時間と能力がないため、ブランドの集合である「カテゴリー」単位で評価を行うとされている [5-7]。カテゴリー単位で意思決定を行うことを「カテゴリーベース処理」と呼ぶが、カテゴリーベース処理を仮定した選択行動モデルは、とくに FPG 市場においては、ブランド単位の選択行動モデルよりも妥当であると考えられる。そこで本研究では、ブランド単位の選択行動モデルではなく、カテゴリー単位の選択

かつまた そうたろう, ういらわん どに だはな
大阪大学大学院経済学研究科
〒 565-0043 大阪府豊中市待兼山町 1-7
katsumata@econ.osaka-u.ac.jp
dony@econ.osaka-u.ac.jp
なかじま のぞみ
宮城学院女子大学現代ビジネス学部
〒 981-8557 宮城県仙台市青葉区桜ヶ丘 9-1-1
nakajima@mgu.ac.jp
受付 16.7.25 採択 16.11.25

行動モデルを構築し、その有用性を実証的に検討していく。

2. カテゴリーベース処理の理論

2.1 FPG 市場におけるカテゴリーベース処理

消費者行動研究におけるカテゴリーとは、代替案（ブランド）を識別するための認知構造の一つであり、ある共通した「定義的特性」によって分類される代替案（ブランド）の集合と定義される [6, 7]。このカテゴリーに注目し、消費者のカテゴリーベース処理を検討した初期の研究としては [8, 9] がある。カテゴリーベース処理は、ブランドの属性を個別に検討する多属性態度モデルに代表される「属性ベース」のボトムアップ型の処理ではなく、トップダウン型の処理であるとされ、より高速な意思決定が可能であるとされている。高速な意思決定は、とくに FPG 市場の選択において大きな有用性をもつ。たとえばスーパーマーケットにおける製品購入は 1 回の来店で数十品目になり、1 回の来店中に数十の意思決定を行っていることになる。このような環境では、「最適な」ブランド選択を目指す時間のかかる処理ではなく、「満足できる」ブランド選択を目指す高速な処理を行うと考えられる。また、多くのブランドが存在する市場においては、情報過負荷に陥りやすくなるために不要な情報を排除しなければならないが、カテゴリーベースであれば情報処理は簡易になる [10]。先行研究からも、FPG 市場におけるブランド選択は、多くはカテゴリーベース処理によるものであると考えることが妥当といえる [5, 6, 11]。

2.2 カテゴリーの軸と駆動するスキーマ

2.1 節の議論から、FPG 市場において消費者がカテゴリーベース処理で意思決定を行っているケースは多いと考えられるが、消費者がブランドの集合をどのようなカテゴリーに分類しているかという問題を解決しなければならない。消費者は個人ごとに異なるカテゴリー事前知識をもっており、事前知識を駆動させてカテゴリーが認知されるといわれている [6, 7]。この事前知識をスキーマ (schema) と呼び、駆動するスキーマが異なると、消費者が認知するカテゴリーの軸も異なってくると考えられる。たとえば、具体的なブランドでいえば「アサヒ・スーパードライ」は、分類軸として「企業」と「酒税法上の酒類分類」があると考えられる。ここで、「企業スキーマ」が駆動すれば、アサヒの飲料として認識され、「酒類スキーマ」が駆動すれば「ビール」飲料として認識される。どのスキーマが駆動し、どのカテゴリー軸で製品を見ているかは個人

差があるが、一定の傾向があると考えられる [12, 13]。FPG 市場にかかわらず多くの市場において、市場に参入しているメーカーが市場ポジションの異なる定番製品ブランドをもっていることが多く、類似したポジションに競合ブランドを投入するので、製品ラインが社会的共通認識として制度化していくという現象がみられる [14–16]。市場には制度化されて共有されたカテゴリーがいくつか存在し、消費者はある程度共通したカテゴリーの軸をもっていると想定される。

上記の、カテゴリーベース処理や、市場における制度化されたカテゴリーの議論を踏まえ、本研究では、消費者のカテゴリー選択に主眼をおいて分析を進めていく。ブランド単位での分析ではないが、企業にとっても、カテゴリーに注目することは市場戦略上有用であると考えられる。一般に、多くの市場はいくつかのセグメントに分割されており、各企業は自社ブランド同士のカニバリゼーションを避けるために、一つのセグメントにあまり多くのブランドは投入しない。したがって、セグメントごとにカテゴリーが形成されていると考えて分析を進めても、個別ブランド戦略と同等の示唆を得ることができると考えられる。

3. データとモデル

3.1 データと対象市場

本研究で用いたデータは、「経営科学系研究部会連合協議会主催 平成 27 年度データ解析コンペティション」を通じて (株) アイディーズから貸与されたスーパーマーケットの ID 付き POS データである。貸与されたデータには 9 店舗分の購買履歴情報があるが、分析対象店舗としては、この 9 店舗のうち、関東地方の 1 店舗としている。店舗選択は無作為であるが、本研究で提案するモデルは、ほかの店舗でも同様に適用することができる。

対象市場としては、ビール系飲料を選択した。ビール系飲料市場は、「ビール」、「発泡酒」、「新ジャンル (第三のビール)」で構成される酒類市場の一部であり、典型的な FPG 市場であるといえる。また、大手 4 社による寡占市場であり、制度化されたカテゴリーの軸が存在している。これを踏まえて、本研究では次のようなカテゴリーの分類軸を仮定する。

- ・ カテゴリー軸 1: 製品ライン={ ビール, 発泡酒, 新ジャンル }
- ・ カテゴリー軸 2: 企業={ キリン, アサヒ, サントリー, サッポロ, その他 }
- ・ カテゴリー軸 3: 新規性={ 新製品, 既存製品 }

分析対象とする製品はブランド名を基準に分類している。まず、購買履歴データとは別途貸与されたSKU単位の製品リストのデータベースから、ビール系飲料について1,494個のSKUを得て、1,494個のSKUに対して目視でブランドを集約し、250種類に分類している。サイズ違い、4本入り、6本入りなどは同じブランドとして分類しているが、拡張ブランド、期間限定ブランド、キャンペーンパッケージは別ブランドとして分類している。すべてのブランドに対して、いずれかの製品ライン、いずれかの企業が紐付くことになる。なお、分析対象店舗で購入されたブランドは164種類であった。また、新規性の判定は購買履歴から行っている。詳細な方法については、付録1を参照されたい。なお、新しく発売された製品と判定されたブランドは100種類である。

また、分析対象の消費者は、2年間で20回以上ビール系飲料を購入した700人とした。購買数が多ければカテゴリーベース処理を行うと考えられるため[17, 18]、一定回数以上当該市場の製品を購入している消費者を選択している。20回は、2年間の分析期間を考えると、平均して5週間に一度購入していることになる。おおよそ1カ月であり、本研究では、この頻度以上の購買回数があれば継続的な購買があり、カテゴリーベース処理のような簡易な処理によってブランド選択をしていると想定している。分析対象となった700人について、2年間の購買回数は、中央値39回、平均値53.38回、最大195回である。男女比については、男性539人(77%)、女性161人(23%)である。年齢は平均値56.7歳、中央値56歳である。

3.2 モデル

3.1節で定義したカテゴリー軸について、離散選択モデルをもとに実証分析のモデルを構築する。本研究の仮定では、消費者 $i \in \{1, \dots, N\}$ の $t \in \{1, \dots, T_i\}$ 回目の購買機会においては、三つの意思決定があることになる。たとえば、消費者 i が購買機会 t に「サントリー ザ・モルツ」を購入したとき、企業は「サントリー」が選択されたことになり、製品ラインは「ビール」が選択されたことになる。さらに、このブランドを過去購入していなければ「新製品」が選択されたことになる。本研究では、個別のブランドではなく、選択したブランドの「製品ライン」「企業」「新規性」が観測対象となる。したがって、観測される選択行動は以下の三つに分割される。

- ・製品ライン選択： $y_{it1j} \in \{0, 1\}$, $j = 1, \dots, J_1$
- ・企業選択： $y_{it2j} \in \{0, 1\}$, $j = 1, \dots, J_2$

- ・新規性選択： $y_{it3j} \in \{0, 1\}$, $j = 1, 2$

製品ライン j で企業 k が販売している新製品が選択される確率は $\Pr(y_{it1j} = 1 \wedge y_{it2k} = 1 \wedge y_{it31} = 1)$ となる。新製品の判定については、消費者個人にとって未購買のブランドであるかを判定している。詳細な方法については、付録1および図4を参照されたい。また、上記の選択の背景には、潜在変数を仮定する[19, 20]。

$$y_{itdj} = \begin{cases} 1, & \text{if } u_{itdj} = \max_k(u_{itdk}), d = 1, 2, 3 \\ 0, & \text{if } u_{itdj} < \max_k(u_{itdk}) \end{cases} \quad (1)$$

続いて、潜在変数 u_{itdj} を説明する構造を考える。説明変数には、当該軸におけるカテゴリーベース処理が活性化することが想定される要因を組み込んでいく。説明変数は大別して3種類ある。第一は社会要因である。新聞記事とGoogle Trendsの検索数インデックスを用いる。新聞記事(NP_{jt})は日経3紙(日本経済新聞、日経産業新聞、日経MJ)のうち、本文にキーワードが登場した回数(当日を含めた14日間移動平均)をおく。検索数(IS_{jt})はGoogle Trendsの週次検索数(当日以前直近の検索数)。ただし、企業「その他」については全期間中「0」とおく。いずれも標準化して利用する。第二は価格(PR_{jt})である。日別でブランドごとに100ml当たりの販売価格を計算し、カテゴリーごとに平均値をとる。また、新規性意思決定については、全ブランドの平均価格を組み込んでいる。また、購買があったブランドについては、実売価格(100ml当りに換算)を利用している。第三は入手可能ブランド数(AV_{jt})である。前日から1週間以内に販売実績のあるブランドを「入手可能ブランド」とし、カテゴリーごとに総和をとる。新規性意思決定については、入手可能な全ブランド数を説明変数として組み込んでいる。また、入手可能ブランド数については、あまりに多すぎると効用が低下する可能性があるため、二乗項もあわせて組み込んで、影響関係を検討する。 $X_{jtd} = \{I_{j\cdot}, NP_{jtd}, IS_{jtd}, PR_{jtd}, AV_{jtd}, AV_{jtd}^2\}$, $d = 1, 2, 3$ である。ただし、 I は切片にあたる単位行列で、 $I_{j\cdot}$ はその第 j 行である。加えて、時間変化を捉えるパラメータ μ_{it} も組み込む。 μ_{it} からは個人の動的な選好の変化を読み取ることができる。以上の説明変数とパラメータを組み込み、各軸の各代替案の潜在変数を推定するモデルを以下のように定義する。識別性を満たすために、潜在変数、説明変数ともにベース代替案との差をとっており、 $u_{itd}^*, X_{itd}^*, \mu_{itd}^*, e_{itd}^*$, $d = 1, 2, 3$ はベースとなる代替案から差をとった変数である([21]を参

照のこと)。ベースとなる代替案は、製品ライン軸では「新ジャンル」、企業軸では「その他」、新規性軸では「既知」をおいている。

$$\begin{aligned}
 u_{it}^* &= X_{it}^* \beta_i + \mu_{it}^* + \varepsilon_{it}^*, \varepsilon_{it}^* \sim N(0, \Sigma), \\
 X_{it}^* &= \begin{pmatrix} X_{it1}^* & O & O \\ O & X_{it2}^* & O \\ O & O & X_{it3}^* \end{pmatrix}, \\
 u_{it}^* &= \begin{pmatrix} u_{it1}^* \\ u_{it2}^* \\ u_{it3}^* \end{pmatrix}, \beta_i = \begin{pmatrix} \beta_{i1} \\ \beta_{i2} \\ \beta_{i3} \end{pmatrix}, \\
 \mu_{it}^* &= \begin{pmatrix} \mu_{it1}^* \\ \mu_{it2}^* \\ \mu_{it3}^* \end{pmatrix}, \varepsilon_{it}^* = \begin{pmatrix} \varepsilon_{it1}^* \\ \varepsilon_{it2}^* \\ \varepsilon_{it3}^* \end{pmatrix}
 \end{aligned} \quad (2)$$

個人ごとに推定され、時不変のパラメータ β_i については、次のような階層構造を仮定する。ただし、 w_i は消費者個人特性の説明変数である。本研究では、 w_i には切片、年齢の対数、性別（女性 = 1）を組み込んでいる。 Γ は行列型のパラメータであり、新聞記事数や入手可能ブランド数などの反応係数の大きさと消費者のデモグラフィック属性との関係を検討することができる。

$$\beta_i = \Gamma w_i + \eta_i, \eta_i \sim N(0, V) \quad (3)$$

また、時間変化するパラメータ μ_{it}^* については、次のような構造を仮定する。

$$\mu_{it}^* = \mu_{(i,t-1)}^* + \nu_{it}, \nu_{it} \sim N(0, \Omega) \quad (4)$$

ただし、識別性を保証するために、 $\mu_{i1}^* = 0$ とおく。

モデルは上記のような、3種類のプロビットモデルが並列する構造をもつ。代替案の数は、製品ライン軸で3項、企業軸で5項、新規性軸で2項であるが、識別性をもつモデルの次元は $3-1+5-1+2-1 = 7$ 次元である。また、上記に示した以外にも、分散共分散行列に識別条件があるが [21]、詳細な条件については、付録を参照されたい。推定はマルコフ連鎖モンテカルロ (MCMC) 法を使っている (e.g., [3, 22])。プロビットモデルがベースになっているため、推定においては [20] および [23] の方法を拡張して活用している。また、時変のパラメータについては、状態空間モデルとなっているため、カルマンフィルタと固定区間平滑化を行っている。詳細な推定の手続きについては [24, 25] に詳しい。

3.3 アンカー・カテゴリーと選択確率の集中度

消費者がどのような軸でブランドを分類しているのか、すなわち、どの軸のスキーマが主として駆動しているのか、本研究では、3.2節で定義した選択確率から主として駆動するスキーマを検討していく。消費者が第一に検討するカテゴリーの軸は、[6]においては「アンカー・カテゴリー」と呼ばれている。これは [12, 26] によって検討されている階層のカテゴリー構造において最上位にあるカテゴリー分類軸とも捉えることができ、消費者の意思決定の大きな基準となるカテゴリー軸である。本節では、この「アンカー・カテゴリー」を検討する基準について説明する。

3.2節で提示したモデルを使うと、各軸について、代替案の選択確率が計算される。たとえば「製品ライン軸」であれば、「ビール」「発泡酒」「新ジャンル」の選択確率が、総和が1になるようにそれぞれ与えられる。同時に、「企業軸」では、「キリン」「アサヒ」「サントリー」「サッポロ」「その他」の選択確率が、これも総和が1になるようにそれぞれ与えられる。ここで、各軸について、特定の代替案に選択確率が集中する程度、選択確率集中度を定義する。まず、カテゴリー分類軸を $d = 1, 2, 3$ とおき、カテゴリー軸 d における消費者 i の選択確率を $p_{id} = (p_{id1}, \dots, p_{idJ_d})$ とおく。実際のモデルでは p_{itdj} だが、本節では t を省略して説明する。得られた選択確率を大きい順にソートした列を \tilde{p}_{id} とおく。すなわち、 $\tilde{p}_{idj} \geq \tilde{p}_{id,j+1}$, $j = 1, \dots, J_d - 1$, $\tilde{p}_{idJ_d} = \min_j(\tilde{p}_{idj})$ となる。 \tilde{p}_{idj} から、選択確率の集中度を上位 j 個の代替案まで総和した累積選択確率を $\tilde{P}_{idj} = \sum_{j'=1}^j \tilde{p}_{idj'}$ とおくと、消費者 i の軸 d に対する選択確率集中度 CR_{id} は次の式 (5) から得ることができる。

$$CR_{id} = \sum_{j=1}^{J_d} \left(\tilde{P}_{idj} - \frac{j}{J_d} \right) / \left(\frac{J_d - 1}{2} \right) \quad (5)$$

選択確率集中度 CR_{id} は、軸 d において特定の代替案に選択が集中する程度と解釈することができる。また、 $CR_{id} \in [0, 1]$ の値をとり、0で完全均等、1に近いほど集中度が高いといえる。選択確率集中度 CR_{id} のイメージとしては、図1にあるように、横軸にソートした代替案をとり、縦軸に累積選択確率をとった図の対角線と累積選択確率 \tilde{P}_{idj} が描く曲線との間の面積を、最小0、最大1に変換したものである。

ただし、この CR_{id} は市場全体のシェアの偏りから影響を受けると考えられる。そこで、異なるカテゴリー分類軸の CR を比較するときは、カテゴリー分類軸 d

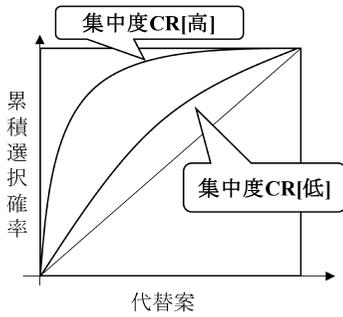


図1 CRの視覚的イメージ

表1 市場全体のシェア

企業	シェア	累積シェア
キリン	0.411	0.411
アサヒ	0.252	0.663
サッポロ	0.142	0.805
サントリー	0.139	0.945
その他	0.055	1.000
製品ライン	シェア	累積シェア
新ジャンル	0.476	0.476
ビール	0.366	0.842
発泡酒	0.158	1.000

における市場全体の集中度である MCR によって調整した CR_{id}^* を比較に用いる。ただし、 CR_{id}^* は 1 より大きい値を取りうることに注意されたい。

$$CR_{id}^* = \frac{CR_{id}}{MCR_d} \quad (6)$$

表1はとくに企業、製品ラインについて、分析対象の店舗データを用いて計算したシェア p_j 、累積シェア \tilde{P}_j である。ここから CR と同様に式(5)を使って市場全体の集中度 MCR を計算すると、それぞれ 0.412、0.318 となる。また、同様の方法で新製品の MCR を算出すると、0.745 となった。アンカー・カテゴリーの考察においては、この MCR で調整した CR^* を用いる。

選択確率集中度について、理論的な背景は以下のとおりである。まず、消費者はカテゴリーベース処理によって、代替案を部分集合単位で特徴づけ、評価していることになる [7]。カテゴリー化は選択を高速かつ容易にするための方略であるため、当該カテゴリー軸のスキーマが活発に駆動していれば、各代替案（正確にはブランドの部分集合になる）について「積極的に選好される代替案」と「忌避される代替案」が明確に分離していることになる。すなわち、いくつかの軸の

中でこの「選好される代替案」と「忌避される代替案」の差が最も大きい軸は、代替案の判別が明確であるといえ、当該軸におけるスキーマが最も活発に駆動しており、ひいては当該軸がアンカー・カテゴリーであるといえる。たとえば、ある消費者について、「製品ライン」軸で「ビール」の選択確率がほかの代替案（「発泡酒」、「新ジャンル」）と比較して極めて高く、かつ「企業」軸ではすべての企業の製品がほぼ等確率で選択されるような状況のとき、その消費者は「製品ライン」軸のスキーマが活発に駆動し、代替案の判別を行っているといえる。これをモデルから導くと、ある軸について選択確率の偏りを数値で捉える CR を計算することで、その軸のスキーマがどの程度駆動しているかを比較することができるといえる。とくに、市場全体の偏りを調整した CR^* を用いることで、より精緻な比較ができると期待できる。

3.4 予測とモデル比較

3.2 節で構築したカテゴリーベース処理の選択モデルについて、本研究では二つのモデルを並行して推定し、モデルの全体的な適合を検討する。まず、モデル1は「動的モデル」であり、選好の時間的な変化を仮定したモデルである。このモデルでは3.2 節で定義した μ_{it}^* を未知パラメータとして推定している。モデル2は「静的モデル」であり、選好の時間的な変化を仮定しないモデルである。このモデルでは動的モデルから μ_{it}^* を外したものである。比較基準としては、予測能力を検討する。各消費者について、最後の購買を推定サンプルから除外してこれを予測する。推定について、前節までで言及したようにMCMC法を用いており、繰り返し回数は10,000回であり、そのうちBurn-in 5,000回を除いた5,000回をサンプルとして取得している。

4. 結果

4.1 予測能力

まずはモデルの予測能力を検討する。表2について、ヒットレートは、カテゴリーごとに予測を行っている。完全予測は「製品ライン」「企業」「新規性」の3点すべてについて予測が当たっている消費者の割合である。また、ROCスコアは、ROC曲線を描画したときの曲線下部の面積であり、完全ランダム予測では0.5が期待され、0.5よりも大きければ予測能力があるとされる。最大で1をとり、1に近いほど予測能力が高いと判定できる [27]。表2を見ると、動的モデル、静的モデルともに、高い予測能力があることが見てとれる。とくにROCスコアは0.5を大きく上回っており、予測

モデルとしても十分活用できるといえる。二つのモデルの比較においては、どちらかのモデルがすべてにおいて優れているという結果ではないが、同等に予測能力は高いといえる。以後の考察においては、時間変化する要素も検討できる動的モデルを検討対象とする。

4.2 パラメータの推定結果

本節では、全体的な傾向を検討するために、 β_i の事前構造として仮定されているパラメータ Γ を検討対象とする。表3は、得られたパラメータ Γ の推定結果である。斜体で示されている数字は、10%最高事後密度区間 (HPDI) が0を含まなかったもの、また、太字で示されている数字は、5%HPDIが0を含まなかった

表2 予測結果の比較

	ヒットレート	静的モデル	動的モデル
個別	製品ライン	0.810	0.817
	企業	0.671	0.683
	新規性	0.840	0.813
総合	完全予測	0.554	0.583
	2点予測	0.250	0.197
	1点予測	0.159	0.170
	ROCスコア	静的モデル	動的モデル
製品ライン	ビール	0.929	0.921
	発泡酒	0.930	0.924
	新ジャンル	0.915	0.904
企業	キリン	0.866	0.865
	アサヒ	0.865	0.865
	サントリー	0.843	0.823
	サッポロ	0.906	0.875
	その他	0.980	0.948
新規性	新製品購入	0.638	0.641

ものである。特徴ある傾向が出た変数としては、まず製品ラインについて、新ジャンルは若い消費者に選好されていることがわかる (年齢・製品ライン・新ジャンル = 0.144)。また、企業についてはすべて年齢の高い消費者に選好されていることがわかり、年齢の高い消費者は「その他」の企業と比較するとナショナルブランドの製品を選好する傾向が高いといえる (年齢・企業・キリン = 0.199, アサヒ = 0.379, サントリー = 0.260, サッポロ = 0.335)。性別では、とくにサッポロが女性に好まれるようだ (性別・企業・サッポロ = 0.336)。また、価格について、価格と切片を見ると、製品ライン軸、企業軸においては価格の高いものは忌避される傾向にあるが (-0.628, -0.840)、新規性に関しては価格の影響は強くないことがわかる (-0.424)。同様に、価格と年齢を見ると、とくに企業軸において、価格と年齢のパラメータの値が正であり (0.182)、年齢が低い消費者は高価なブランドを避ける傾向があることが示唆される。ほかに傾向がみられた点としては、新聞記事の出稿が増える製品ラインは選択されやすい傾向があることがわかる (0.033)。

4.3 選択確率集中度

3.3節で定義した選択確率の集中度から、各個人の選択確率集中度 (CR) を計算することができる。

表4は、本研究で注目した製品ライン、企業、新規性の各軸について式 (5), (6) から市場シェアを調整した選択確率集中度 CR^* を計算し、個人ごとに順序を付してまとめたものである。製品ライン軸の CR^* がもっとも高かった消費者は全体の71.0%であり、企業軸の CR^* がもっとも高かった消費者は全体の5.9%で

表3 パラメータ Γ の推定結果

	切片			年齢			性別		
	製品ライン	企業	新規性	製品ライン	企業	新規性	製品ライン	企業	新規性
ビール	0			0			0		
発泡酒	-0.820			0.094			-0.064		
新ジャンル	0.109			0.144			0.289		
キリン		-0.788			0.199			-0.114	
アサヒ		-0.171			0.379			0.156	
サントリー		-1.152			0.260			0.216	
サッポロ		-0.963			0.335			0.336	
その他		0			0			0	
新製品			0.123			0.527			0.168
既存製品			0			0			0
新聞記事	0.033	-0.007	-0.003	-0.007	0.032	0.000	0.009	-0.011	0.014
検索数	0.085	0.342	0.137	0.326	-0.224	-0.043	0.094	-0.263	-0.066
入手可能ブランド数	0.362	0.388	0.067	0.152	0.157	-0.357	-0.254	-0.218	0.284
入手可能 (2乗)	-0.150	-0.547	1.010	0.549	-0.063	0.472	0.366	-0.282	-0.156
価格	-0.628	-0.840	-0.424	0.064	0.182	-0.110	0.157	0.013	-0.068

表 4 CR* の順位

CR* 順位	製品ライン	企業	新規性
第 1 位	497	41	162
	71.0%	5.9%	23.1%
第 2 位	88	224	388
	12.6%	32.0%	55.4%
第 3 位	115	435	150
	16.4%	62.1%	21.4%
計	700	700	700

あることがここから見てとれる。この順位は、[26] によって議論されている階層のカテゴリー構造としても捉えることができる。

第 1 位の軸がアンカー・カテゴリーとなり、階層のカテゴリー構造の最上位に位置し、第 2 位の軸が第二の階層、第三の軸がその下の階層にあると考えることもできる。第 2 位以下の軸は、[7] によって「サブカテゴリー」とも呼ばれている軸ともいえる。すなわち、個人によってどの軸をアンカー・カテゴリーとして、どの軸をサブカテゴリーと捉えているのかが異なるという示唆を得ることができる。

また、図 2 はある個人（消費者 id 33 および 27）の「製品ライン」と「企業」の CR について、購買機会ごとの変動を捉えたものである。1 と振られている位置が、初回の購買機会における CR の値である。45 度線よりも右下にあれば「製品ライン軸」の CR が相対的に高く、製品ラインスキーマが相対的に活発に駆動しているといえる。左上にあれば、「企業」軸の CR が相対的に高く、企業スキーマが活発に駆動しているといえる。図 2 から、消費者の駆動するスキーマが時間変化している様子を見ることができる。図 2 上にある消費者 33 は、観測期間の初期は企業スキーマ駆動したり製品ラインスキーマが駆動したり定まらないが、観測期間の後期には製品ラインスキーマが駆動しているといえる。また、図 2 下の消費者 27 は、観測期間全般にわたって製品ラインスキーマが駆動しているといえる。CR から、消費者が主として検討しているカテゴリーの軸を推測することができ、消費者個人がもつ市場構造や競合関係についての示唆を得ることができる。

5. 議論と結論

5.1 研究の貢献

まずは本研究の貢献として、実務的貢献を挙げる。まず、流通企業の市場戦略について、消費者が検討しているカテゴリーは、棚割りの方法を再検討するための判断基準とすることができる。多くの店舗では製品

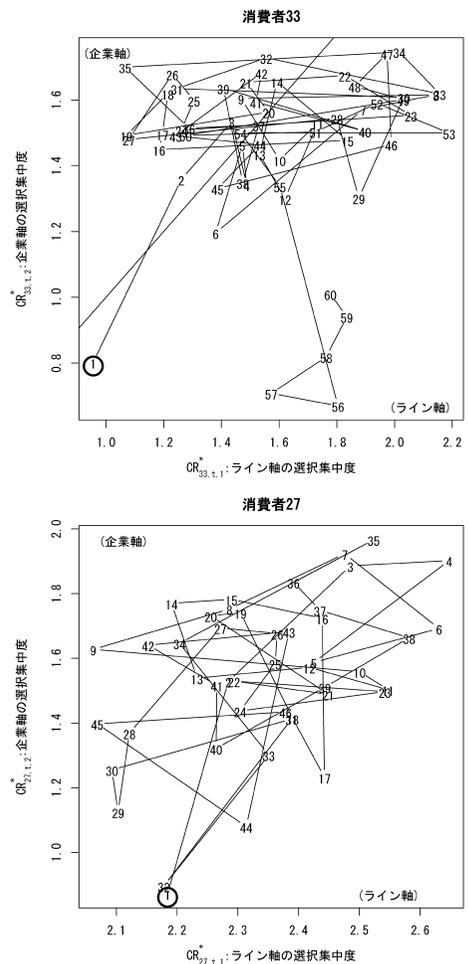


図 2 個人ごとの選択確率集中度の推移

ライン基準で棚割りが決まっているが、製品ライン基準の棚割りでは、企業スキーマで選択している消費者にとっては、比較検討対象とする製品が分散しており、処理に負荷がかかっているといえる。本研究で提案したモデルを用いることで、消費者のアンカー・カテゴリーが何かを知ることができ、それが全体の何割くらい存在しているのかを購買履歴データから知ることができ、棚割りに活用することができる。たとえば、列に企業、行に製品ラインを配置するマトリックス型の棚割りなどが考えられる。

次に、選択行動研究に対する貢献を挙げる。本研究では、ブランド単位ではなく、カテゴリー単位で評価を行う消費者行動モデルの有用性を実証した。得られた結果から、FPG 市場における実用性だけでなく、予測能力も高いことが示された。また、消費者行動研究におけるカテゴリー理論に対して、カテゴリーベース処理において駆動するスキーマについて、実験デー

タではなく行動データから推測する枠組みを整備したことも貢献である。

6. 課題

最後に今後の課題を挙げたい。今後の課題としては、大別して三つが挙げられる。第一は、カテゴリーベース処理に関する議論の精緻化である。本研究では、消費者はカテゴリーベース処理によって選択を行うと仮定しているが、消費者の中にはピースミール処理（ブランドの属性を個別に検討する詳細な情報処理）を行っている者もいると考えられる。カテゴリーベース処理とピースミール処理のどちらを行っているかは、ID-POSデータからは観測が困難であるが、さらに精緻な考察が必要であろう。本研究では、一つの基準として一定回数以上（20回）購買した消費者を対象としているが、何回以上購買すればカテゴリーベース処理となるのか、明確な基準はなく、今後の検討が必要である。

第二は、説明変数の再考である。一つは、より購買行動を高い精度で説明する説明変数を探索することが重要といえる。本研究で計算する選択確率集中度 CR および市場シェアで調整を行った CR^* は、モデルの適合がよければ高い精度で消費者のアンカー・カテゴリーを推定することができるが、適合の低いモデルでは、適切な結果が得られない恐れがある。今後さらなる研究が必要である。ビール系飲料の購買行動では、本研究で組み込んだ以外の要因も大きな影響を与えている可能性がある。たとえば、ビール系飲料市場では季節限定ブランドがあり、季節による需要の変動も大きい。そのため、季節性の要因を組み込むことも必要であろう。本研究で説明変数として用いたインターネット検索数は季節の変動があるため、ある程度季節変動を吸収していると考えられるが、ほかの要因も検討すべきであろう。第三は、分類軸の再考である。本研究で提案したモデルにおいて、分類軸は分析者がモデルに与えなければならない。適切な軸、あるいは新しい軸でブランドを分類することができれば、企業の大きな競争力になる。軸の抽出法については、[26]でも具体的な方法が指示されており、質問票調査や消費者実験の研究からも有用な知見を得ることができるだろう。たとえば本研究で分析対象としたビール系飲料においても、本研究で挙げた以外に、「サイズ」スキーマ、「箱」スキーマで選択をしている消費者もいる可能性が高い。軸の追加は、本研究での提案モデルを拡張することで容易に検討が可能であるが、今後さらなる研究によって適切なモデルを検討していくことが望まれる。

謝辞 本研究の分析データは経営科学系研究部会連合協議会および株式会社アイデイズ様より提供いただきました。また、匿名のレフェリー2名からは大変有用なコメントをいただきました。この場をお借りして御礼を申し上げます。

参考文献

- [1] D. Talukdar, D. K. Gauri and D. Grewal, "An empirical analysis of the extreme cherry picking behavior of consumers in the frequently purchased goods market," *Journal of Retailing*, **86**(4), pp. 336–354, 2010.
- [2] 勝又壮太郎, 西本章宏, 『競争を味方につけるマーケティング』, 有斐閣, 2016.
- [3] P. E. Rossi, G. Allenby and R. McCulloch, *Bayesian Statistics and Marketing*, Wiley, 2005.
- [4] 勝又壮太郎, 阿部誠, "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," *オペレーションズ・リサーチ: 経営の科学*, **57**(2), pp. 53–62, 2012.
- [5] J. Meyers-Levy and A. Tybout, "Schema congruity as a basis for product evaluation," *Journal of Consumer Research*, **16**, pp. 39–54, 1989.
- [6] 新倉貴士, 『消費者の認知世界』, 千倉書房, 2005.
- [7] 西本章宏, 『外部マーケティング資源としての消費者行動』, 有斐閣, 2015.
- [8] M. Sujan, "Consumer knowledge: Effects on evaluation strategies mediating consumer judgments," *Journal of Consumer Research*, **12**, pp. 31–46, 1985.
- [9] M. Sujan and J. R. Bettman, "The effects of brand positioning strategies on consumers' brand and category perceptions: Some insights from schema research," *Journal of Marketing Research*, **26**, pp. 454–467, 1989.
- [10] J. P. Peter and J. C. Olson, *Consumer Behavior*, 9th edition, McGraw-Hill, 2010.
- [11] B. Loken and J. Ward, "Measures of the attribute structure underlying product typicality," *Advances in Consumer Research*, **14**, pp. 22–26, 1987.
- [12] 清水聰, 『戦略的消費者行動論』, 千倉書房, 2006.
- [13] L. A. Peracchio and A. M. Tybout, "The moderating role of prior knowledge in schema-based product evaluation," *Journal of Consumer Research*, **23**, pp. 177–192, 1996.
- [14] A. Humphreys, "Megamarketing: The creation of markets as a social process," *Journal of Marketing*, **74**(2), pp. 1–19, 2010.
- [15] J. A. Rosa, J. F. Porac, J. Runser-Spanjol and M. S. Saxon, "Sociocognitive dynamics in a product market," *Journal of Marketing*, **63**(Special Issue), pp. 64–77, 1999.
- [16] J. A. Rosa and J. Spanjol, "Micro-level product-market dynamics: Shared knowledge and its relationship to market development," *Journal of the Academy of Marketing Science*, **33**(2), pp. 197–216, 2005.
- [17] J. W. Alba and J. W. Hutchinson, "Dimensions of consumer expertise," *Journal of Consumer Research*, **13**, pp. 411–454, 1987.
- [18] M. Sujan and C. Dekleva, "Product categorization and inference making: Some implications for comparative advertising," *Journal of Consumer Research*, **14**, pp. 372–378, 1987.
- [19] J. H. Albert and S. Chib, "Bayesian analysis of bi-

nary and polychotomous response data,” *Journal of the American Statistical Association*, **88**, pp. 669–679, 1993.

- [20] R. McCulloch and P. E. Rossi, “An exact likelihood analysis of the multinomial probit model,” *Journal of Econometrics*, **64**, pp. 207–240, 1994.
- [21] K. E. Train, *Discrete Choice Model with Simulation*, 2nd edition, Cambridge University Press, 2009.
- [22] M. H. Chen, Q. M. Shao and J. G. Ibrahim, *Monte Carlo Methods in Bayesian Computation*, Springer, 2000.
- [23] Y. D. Edwards and G. M. Allenby, “Multivariate analysis of multiple response data,” *Journal of Marketing Research*, **40**(3), pp. 321–334, 2003.
- [24] J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*, Oxford University Press, 2003.
- [25] 片山徹, 『応用カルマンフィルタ 新版』, 朝倉書店, 2000.
- [26] 清水聰, “知識カテゴリーの実証研究,” *消費者行動研究*, **10**(1–2), pp. 1–15, 2004.
- [27] R. C. Blattberg, B-D. Kim and S. A. Neslin, *Database Marketing: Analyzing and Managing Customers*, Springer, 2008.

付録

1. 新製品の判定

本研究では、新製品の判定は「消費者にとって未購買の製品」であることを基準とする。判定の手続きとしては、まず、ブランドごとに分析期間中に発売されたブランドである確率を推定し、「分析期間中に発売されたブランド」が消費者個人にとって初回の購買であれば、その消費者が「新製品を購入した」と判定する。期間中に発売された製品である確率は以下のモデルから推定する。あるブランド j について、一日の購買数が時間不変パラメータ λ_j をもつポアソン分布に従っていると仮定する。期間中の合計購買数が S_j であるとする。また、期間中最後の購買が観測された日を T_j とおく。 S_j は期間中の購買数の和なので、一日当たりの購買数は $s_j = S_j/T_j$ となる。また、合計購買数は $S_j \sim \text{Poisson}(T_j \lambda_j)$ である。このブランドの購買が初めて観測されるのが τ_j であるとする。ここから、このブランドが時間 1 から入手可能であり、 τ_j 時点まで購買が観測されない確率として以下を得る。

$$\prod_{t=1}^{\tau_j-1} \Pr(s_t = 0) = \prod_{t=1}^{\tau_j-1} \frac{\lambda_j^0 e^{-\lambda_j}}{0!} = e^{-(\tau_j-1)\lambda_j} \quad (7)$$

ここで、最尤法の解は $\lambda_j = s_j$ であり、 s_j を代入して確率を求めると $q_j = e^{-(\tau_j-1)s_j}$ を得る。この確率は「既存率」であり、「 $1 - q_j$ 」が期間中に発売された製品で

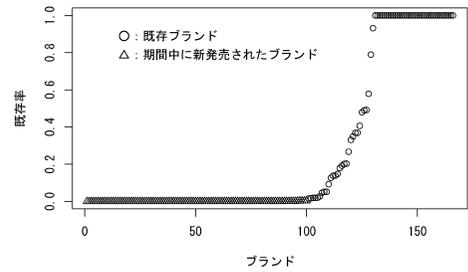


図3 新製品の判定 (既存率)

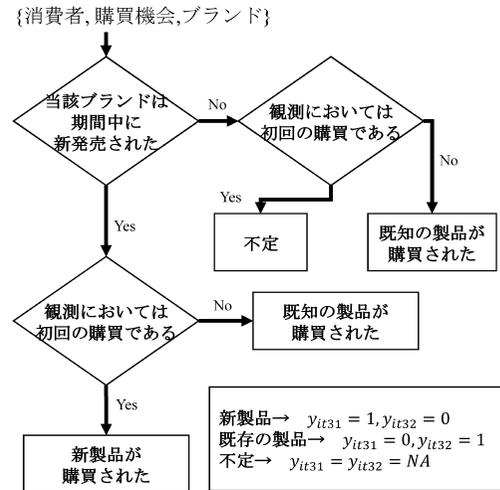


図4 新製品の判定

ある確率となる。本研究では、判定として、 $q_j < 10^{-2}$ のブランドを期間中に発売された製品であると判定している。図3は、分析対象のブランドの既存率をソートしてプロットしたものである。○で示しているブランドが既存製品と判定したもので、△で示されているブランドが期間中に新発売された製品と判定されたものである。また、消費者個人が新製品を購入したか否かの判定は図4のとおりである。

なお、本研究で分析対象としている700人については、分析から除外した8店舗での購買は観測されなかったため、これらの店舗で先に購買しているケースはないといえるが、データの得られていないほかの店舗で先に新製品を購入している可能性もある。これについては観測できないため対応は難しいが、今後の課題としたい。また、毎年一定期間のみ発売される季節限定ブランドは、本研究の定義では新製品と判定されているものが多いので注意されたい。

2. モデルの設定と推定

本節では、モデルの事前分布と、一部の事後分布に

ついて説明する. 説明を付していないパラメータの事後分布については [3], 状態空間パラメータに関しては [24, 25] を参照されたい.

2.1 事前分布

まず, 各変数の次元について, $\dim(u_{it}^*) = J$, $\dim(\beta_i) = K$, $\dim(\Gamma_j) = L$ である. これより, 分散共分散行列のサイズは, Σ が $J \times J$, V が $K \times K$ となる. 各パラメータの事前分布について, $\Sigma^{-1} \sim \text{Wishart}(s_0, S_0)$, $s_0 = K$, $S_0 = 100I_J$, $\Omega_i = \omega_i I_J$, $\omega_i^{-1} \sim \text{Ga}(\zeta_0/2, Z_0/2)$, $\zeta_0 = Z_0 = J$, $\Gamma \sim N_{J \times L}(G_0, V, P_0)$, $G_0 = O_{J \times L}$, $P_0 = 100I_L$, $V^{-1} \sim \text{Wishart}(r_0, R_0)$, $r_0 = J$, $R_0 = 100J$ とおく.

2.2 事後分布

本節では, とくに u_{itj}^* と Σ の事後分布を示す. まず, u_{itj}^* , $i = 1, \dots, N$, $t = 1, \dots, T_i$, $j = 1, \dots, J$ の事後分布は次の切断正規分布から発生させる

$$u_{itj}^* | \cdot \sim \begin{cases} TN_{(a_1, \infty)}(m_1, \rho_1), & \text{if } y_{itj} = 1 \\ TN_{(-\infty, a_1]}(m_1, \rho_1), & \text{if } y_{itj} = 0 \\ N(m_1, s_1), & \text{if } y_{itj} = NA \end{cases} \quad (8)$$

ここで, $m_1 = \nu_{itj} + \Sigma_{j,-j} \Sigma_{-j,-j}^{-1} (u_{it,-j}^* - \nu_{it,-j})$, $\rho_1 = \Sigma_{jj} - \Sigma_{j,-j} \Sigma_{-j,-j}^{-1} \Sigma_{-j,j}$, $\nu_{it} = X_{it}^* \beta_i + \mu_{it}^*$ で

ある. 切断正規分布の閾値 a_1 については以下のとおり, 要素によって異なる

$$a_1 = \max(u_{it,2}^*, 0), \text{ if } j = 1, \quad (9)$$

$$a_1 = \max(u_{it,1}^*, 0), \text{ if } j = 2, \quad (10)$$

$$a_1 = \max(u_{it,4;6}^*, 0), \text{ if } j = 3, \quad (11)$$

$$a_1 = \max(u_{it,3}^*, u_{it,5;6}^*, 0), \text{ if } j = 4, \quad (12)$$

$$a_1 = \max(u_{it,3;4}^*, u_{it,6}^*, 0), \text{ if } j = 5, \quad (13)$$

$$a_1 = \max(u_{it,3;5}^*, 0), \text{ if } j = 6, \quad (14)$$

$$a_1 = 0, \text{ if } j = 7. \quad (15)$$

Σ の事後分布については, [23] あるいは [20] の方法を踏襲して, 識別条件を満たさないパラメータ発生させて事後的に調整する方法をとる. まず識別条件を満足しないサンプル $\tilde{\Sigma}$ を得る.

$$\tilde{\Sigma}^{-1} | \cdot \sim \text{Wishart}(s_1, S_1^{-1}) \quad (16)$$

ただし, $s_1 = \sum_i T_i + s_0$, $S_1 = \sum_{i=1}^N \sum_{t=1}^{T_i} (u_{it}^* - X_{it}^* \beta_i - \mu_{it}^*)' (u_{it}^* - X_{it}^* \beta_i - \mu_{it}^*) + S_0^{-1}$ である.

識別条件は $\Sigma_{11} = \Sigma_{33} = \Sigma_{77} = 1$ であるため, 収集した $\tilde{\Sigma}$ から $\Sigma = Q \tilde{\Sigma} Q$ を得る. ただし, $Q = \text{diag}(\tilde{\Sigma}_{11}^{-0.5}, 1, \tilde{\Sigma}_{33}^{-0.5}, 1, 1, 1, \tilde{\Sigma}_{77}^{-0.5})$ である.