

統計モデルによる消費者理解の可能性

佐藤 忠彦[†]

(受付 2017 年 10 月 30 日；改訂 2018 年 3 月 11 日；採択 4 月 6 日)

要 旨

統計モデルの高度活用は、サービス科学を進展させるための必須のツールであり、今後さらに重要度の増すアプローチである。本稿は、サービスを高度化する際に必須の「消費者理解を深めるには」という観点で統計モデル(特にベイジアンモデル)をどのように活用すべきか?に対する提言を目的とし、周辺研究を整理し、既存研究の紹介を基本として整理した。事例で紹介した2つの研究は、マーケティングを題材としたものであるが、サービス研究に対しても十分な示唆を供するものである。

キーワード：ベイジアンモデリング、消費者異質性、時間的異質性、潜在変数。

1. はじめに

本論は、統計モデルを消費者理解にどのように使うかを議論することが主たる目的である。2つの既発表の事例のエッセンスを紹介し、統計モデルが消費者理解に活用できる可能性を理解してもらいたいと考えている。

現代の需要と供給の関係において、「供給は需要を規定できるのか?」という問いを考えなければならない。この問いは、供給者サイドと需要サイドが出会う場である市場において、供給者サイドと需要サイドのどちらに主導権があるのかという問いに読み替えてもらってもよい。市場がどんどん拡大していく右肩上がりの経済の時代、消費者は人と同じものを持つ、人と同じサービスを楽しむことでその心は満たされていた。しかしながら、現代の消費者は人と少しでも違うものを持ち、他人とは別のサービスを受けることに喜びを感じる。サービスの市場における主たるプレイヤーである消費者のニーズは、細分化、多様化という形式で変化してきている。これらの事実を踏まえると、サービスの市場は「需要が供給を規定する市場」だという認識を持たなければならない。「需要サイドが考える“良いもの”を作れば売れるはず」や「需要サイドが考える“良いサービス”を提供すれば使ってもらえる」といった川下発想に基づく戦略立案が求められているのである。

前段の川下発想に基づく戦略立案のキーは、「消費者を理解すること」である。問題は、「消費者を理解する」の含意は何か?である。「消費者を理解する」ことは、消費者の行動の真のメカニズムを解明することと即断してよいものであろうか? 結論を述べれば、消費者の真の行動メカニズムを解明することは容易ではなく、基本的に不可能である。

企業は、最終的に消費者の真の行動メカニズムの解明を求めているわけではない。もちろん、真の行動メカニズムを解明できるのであればそうしたいが、前段に示したようにそれは基本的に無理であり、その代わりにマーケティング意思決定をする際に最も必要な、時を得た情

[†] 筑波大学 ビジネスサイエンス系：〒112-0012 東京都文京区大塚 3-29-1

- (1) 「消費者を理解する」=「マーケティングを高度化しうる仮説的役割を担う
情報を発見する」

(2) 「消費者を理解する」≠「消費者の真の行動メカニズムを明らかにする」

図1. 「消費者を理解する」の構図.

報を、データ(行動の結果データやアンケートによる態度データ)から抽出したいと考えているのである。換言すれば、企業がデータに基づき行っている「消費者理解のための活動」は、真の構造を解明することを狙っているわけではなく、消費者行動の仮説的役割を担う高次情報の抽出を狙いとしている。以上の検討に基づけば、「消費者を理解する」の含意は何か?に対する答えは、図1のように整理できる。

前段までの議論に基づくと、消費者の真の行動メカニズムを明らかにすることはできないし、そもそもそれを考えることは無意味だといっているように思われてしまうかもしれない。しかし、それは誤解である。消費者の真の行動メカニズムが解明できれば、それにこしたことはないし、当然、可能であるならばそれを実現したい。しかし現実的には、消費者の真の行動メカニズムを解明するのは容易ではないし、実際には何が真の構造なのかさえ分からないため評価もできない。そういった状況でも、「消費者の理解」を深める活動を行わない限り、市場に存在する様々なサービスを高度化できない。

前段までに示した消費者理解をさせるためのキーとなる技術が、ベイジアンモデリングである。以降には、ベイジアンモデリングの概要を示す。なお、 y_t, x_t, ψ は、観測データ、パラメータ、超パラメータをそれぞれ示すものとし、議論を進める。従来の統計モデリングでは少ないパラメータで表現されるモデルが良いモデルとされてきた。所謂、「けちの原理」である。どのようなデータであったとしても平均と分散のみで規定される正規分布から得られたと仮定するようなものである。一方、パラメータ数を増やせば統計モデルの記述能力は向上するが、汎化能力と呼ばれる将来のデータの予測能力が低下する。この問題への対策として、パラメータ x_t についても統計モデル $p(x_t|\psi)$ を想定するのがベイズモデルである。 $p(x_t|\psi)$ をベイズ統計では事前分布と呼ぶ。この事前分布を導入し、ベイズの定理

$$(1.1) \quad p(x_t, \psi|y_t) = \frac{p(y_t|x_t)p(x_t, \psi)}{p(y_t)} \\ \propto \underbrace{p(y_t|x_t)}_{\text{尤度関数}} \underbrace{p(x_t|\psi)}_{x_t \text{の事前分布}} \underbrace{p(\psi)}_{\psi \text{の事前分布}}$$

を用いることで、想定した事前分布 $p(x_t, \psi)$ がどのように修正されるのか、つまりパラメータに関する不確実性がデータによりどの程度修正されたのかを観察するのである。ここで $p(x_t, \psi|y_t)$ を事後分布と呼び、我々の興味はこの事後分布に集約される。なお、データ y_t の発生確率 $p(y_t)$ は x_t, ψ によらない数値になるので、事後分布 $p(x_t, \psi|y_t)$ は(1.1)式右辺2行目に比例する。この仕組みにより、多数のパラメータも安定して推定できるようになり、結果として高い予測能力とデータ記述能力を同時に持つ総合的な統計モデルを構成できる。この一連のモデル化の行為を、通常、ベイジアンモデリングと呼ぶ。(1.1)式を用いる際に t が時点を示すならば x_t は時変パラメータを、個人を示すとすれば個人毎のパラメータとなる。

消費者を理解するためには、「消費者異質性」、「時間的異質性」および「潜在変数」といった観点を意識しなければならない。以降には、それらが何を意味するのかを概説する。

初めに、「消費者異質性」と呼ぶ概念を説明する。現代のマーケティング活動では、マイクロ・マーケティングと呼ぶ「個」に焦点を当てた活動が脚光を浴びている。この活動の前提は、消費者一人一人は異質だという認識であり、一人一人の違いにこそマーケティングを高度化するための情報が含まれており、無視すべきではないと考える点である。消費者は、同一日に同一商品を同一の場所で購買したとしても、人が違えば購買に至るメカニズムに差が生じ、その反応も異なる。 t が消費者を表すとすれば、(1.1)式の x_t が消費者異質性の直接的なモデル表現となる。消費者の行動を規定する説明変数として「商品の価格」を考え、 x がその反応を示すと考えると、消費者の異質性の仮定の下では、 $t \neq t'$ ならば $x_t \neq x_{t'}$ となる。実際には、推定された個人ごとのパラメータ値に大きな差があれば異質性が強いと、逆の場合は異質性は小さいと判断する。

次に、「時間的異質性」と呼ぶ概念を説明する。時間的異質性は、現代のマーケティングにおいて上述の消費者異質性と並び重要な概念である。時間的異質性の仮定の下では、同一の個人であったとしても時点が違えば、その反応には差が生じると考える。 t が時点を表すとすれば、(1.1)式の x_t が時間的異質性の直接的なモデル表現となる。消費者の行動を規定する説明変数として「商品の価格」を考え、 x がその反応を示すと考えると、時間的異質性の仮定の下では、 $t \neq t'$ ならば $x_t \neq x_{t'}$ となる。実際には、推定された時点ごとのパラメータ値に大きな差があれば時間的異質性が強いと、逆の場合は異質性は小さいと判断する。

3つ目として、今日的な視点でマーケティングデータの解析をする際に重要な役割を担う「潜在変数」を説明する。現在、企業には消費者の行動の結果を示すビッグデータが多量に蓄積されている。そのためややもすると、「ビッグデータに含まれる観測変数間の関係性を明らかにすれば消費者の行動は解明できる」といった考えが出現している。しかし、消費者の行動は顕在変数間の関係性の抽出のみで解明できるものではなく、「消費者の行動に至る要因は全てデータとして観測できるわけではない」という認識が不可欠である。拠り所をビッグデータにおくだけでなく、消費者行動理論から経験や勘に至るまで、様々な情報を上手に使う、前述の意味での「消費者の理解」を実現しなければならない。問題は、消費者の行動に影響する潜在変数は何か？ということである。特定の商品の購買を想起した場合は、「家庭内在庫量」、「ブランドロイヤリティ」、「参照価格」などが行動に影響する潜在変数になる。このように具体的な行動を想定すればそれに影響しそうな潜在変数の候補は様々想定できる。もちろん、ここで例示した変数以外にも商品購買に影響する可能性を有する潜在変数は数多くある。しかし、行動に影響する潜在変数は、対象とする消費者の行動が何かによって変化するのは明らかである。実際にモデル化する場合は具体的に示すことになるが、一般的に「消費者の嗜好」、「消費者の好き・嫌いや興味がある・ないといった態度」、「消費者が直面する状況」、「消費者のこれまでの経験」と「消費者の将来を予測する能力」などが消費者の行動に影響する潜在変数の抽象的な説明になる。消費者異質性や時間的異質性に加えて、潜在変数(潜在構造)も考慮し消費者の行動を評価できて初めて、実フィールドで高度に活用可能な情報を手にしうる。これまでマーケティング分野では、本項で説明したような個人ごと、あるいは時刻ごとのような細かい潜在変数を意識することは少なかった。これは、モデルの複雑化に対応可能なデータがなかったのが主たる理由である。しかし、ビッグデータの蓄積が進んだ現在、行動の背後に存在する潜在変数は積極的に考慮していかなければならない。もう少し言えば、潜在変数を考慮しビッグデータの活用を進めない限り、「消費者の理解」は進まないのである。

本節の最後に本特集号のテーマであるサービス研究と以降の事例で取り上げるマーケティング研究の共通点と相違点を提示しておくことにする。サービス研究は、個々のサービスを総合的にデザインするうえで必要な知識と思考の提供を目指している。日本学術会議経営学委員会・総合工学委員会合同サービス学分会(2017)によると、サービス研究には次の5つの領域

が存在する。一つ目は、「サービスマーケティング」であり、共創される価値の解明とその価値のコミュニケーションとデリバリーをデザインする。二つ目は、「サービスオペレーション」であり、価値を具現化し、サービスの仕様・意匠・設計から提供方法を企画・運営する。三つ目は、「サービスマネジメント」であり、サービスを事業体として持続的かつ効果的・効率的に運営するための資源配分と組織内外の体制を設計する。これら3つの領域は、経営学や工学等の領域で個別に教育・研究されてきた。近年、統計を含む情報学の技術を用いて効率的に観察・分析・適用を行う「サービスサイエンス」も重要な領域となっており、これが四つ目の領域となる。最後の五つ目は、ICTやIoTの進展に伴い、個々のサービスを有機的に連携させることで、社会全体により高水準で複合的な価値の提供を目指す「サービスエコシステム」である。それらのサービスの領域では、サービスの提供サイドと需要サイド間とで価値を共創する関係をサービスの一単位(医療サービス、小売りサービス、WEBサービス等)と考え、それぞれサービス全体を対象として研究されることが多い。いずれのサービスでもその高度化は、提供者と需要者間で価値共創を促進することで実現できると考えられている。しかし、サービス提供者が需要サイドの態度、行動を的確に把握できなければ、そもそも共創を促進することはできず、結果需要サイドに支持されるサービスは構築できない。マーケティング研究は、サービス研究のようにある特定サービス全体を対象とすることは少ないが、サービスにおける上述の「提供者による需要サイドの態度、行動の把握」は、やはり最も重要な課題であり、ほとんどのマーケティング研究のテーマとなっている。その意味で、本稿のマーケティング研究の事例は、部分的であるにしる、サービスの根幹部分に対する重要な知見を供するアプローチとして活用できる。重要な点なので言及しておく。

本稿の構成は以下のとおりである。2節には、当該分野における(特に)ペジアンモデリングによる研究事例を、3節には、本稿の目的に合致する既発表の研究事例のエッセンスをそれぞれ示す。4節は、まとめと課題である。

2. 先行研究

本節では、基本的にマーケティング研究を対象とし1節に示した「消費者異質性」、「時間的異質性」および「潜在変数」の観点を含む研究を概観する。

2.1 消費者異質性を含む研究

消費者異質性は、マーケティングにおける重要概念であり、当該概念を取り込んだ研究の数は膨大である。Allenby and Rossi (1998)が指摘したように、マーケティングにおける意思決定において、消費者選好の分布はその中心的な役割を果たす。現代のマーケティング研究では、消費者異質性のモデリングが最重要の課題なのである。隣接分野である計量経済学分野では、異質性を局外母数として捉え、積分消去することにより興味の対象外として取り扱うことは対照的である。Allenby and Rossi (1998)と同一著者により整理されたRossi and Allenby (2003)は、マーケティング・サイエンス領域における階層ベイズモデルの有用性を示した先駆的な論文である。その後マーケティング分野で階層ベイズモデルを用いた研究が増大することになる。マーケティング分野における最重要モデルの1つである離散選択モデルを用いて、消費者異質性を取り扱った研究として、Rossi and Allenby (2003), Terui and Dahana (2006), Terui and Ban (2008), 井上 (2010), Terui et al. (2011), 山田・佐藤 (2012), 日高・佐藤 (2016)等がある。また、ポアソン回帰や負の二項回帰などを含む回帰モデルによって消費者の異質性を取り扱ったものとしては、Manchanda, and Chintagunta (2004), Manchanda et al. (2004), 宮津・佐藤 (2015), 山田・佐藤 (2016)等がある。これらの研究は、全てマルコフ連鎖モンテカ

ルロ法(MCMC法)により推定されている。その技術的詳細に関しては、Chib and Greenberg (1996), Geman, S. and Geman, D. (1984), Hastings (1970)等を参照してほしい。その他、階層ベイズを用いた消費者異質性モデリングの方法論や応用事例に関しては、Rossi et al. (2005), 照井 (2008), 佐藤・樋口 (2013)に整理されている。

2.2 時間的異質性を含む研究

時間的異質性を取り扱ったマーケティング研究は、2.1項に示した消費者異質性を取り扱った研究に比べて極端に少ない。その理由は二つある。一つ目は、マーケティング研究者の動的消費者行動に関する興味の低さであり、二つ目はそもそも動的消費者行動に活用可能な時系列データが少なかったことである。しかしながら、近年のIT技術の革新に伴い粒度の細かい時系列データの蓄積が進み、消費者の動的行動特性の解明にも脚光が当たるようになった。マーケティング分野でなされる時間的異質性を取り扱う研究は、基本的に当該分野で市場反応分析と呼ばれる範疇にカテゴライズされるものが多い。Naik et al. (1998), Kondo and Kitagawa (2000), 山口 他 (2004), 佐藤・樋口 (2008a), Bass et al. (2007)といった研究は、目的変数に売上等の集計量を採用する集計型の動的市場反応を取り扱ったものである。これらのモデルは、基本的に線形・ガウス型の状態空間モデルの枠組みでモデル表現されており、その推定にはカルマンフィルタと最尤法が採用されている。一方でBruce (2008), 佐藤・樋口 (2008b, 2009), 本橋・樋口 (2013), 本橋 他 (2012)は、目的変数に来店やブランド選択等を採用する非集計型の動的市場反応を取り扱った研究である。観測モデルが離散選択型の非ガウス型モデルになるため、線形・ガウス型の状態空間モデルでは表現できず、一般状態空間モデルの枠組みでモデル化され、モデルの推定は、Bruce (2008)ではMCMC + 粒子フィルタ、佐藤・樋口 (2008b, 2009), 本橋・樋口 (2013), 本橋 他 (2012)では粒子フィルタと最尤法が用いられている。集計型の動的市場反応モデルは、その推定が容易にできるという利点もあり、当該コミュニティでも研究が増えつつあるが、非集計型の動的市場反応モデルは、モデル推定に様々な困難が生じるため、まだ発展途上である。しかしながら、いずれの枠組みであったとしても時間的異質性を捉えようとするマーケティング研究は、その仮定の自然さから今後研究が進められなければならない領域となっている。佐藤・樋口 (2013)には、マーケティングにおける状態空間モデルの理論と応用事例が解説されている。

線形・ガウス型・状態空間モデルの具体的なモデリング方法やその周辺技術については、北川 (2005)に詳しい。非線形性や非ガウス分布を内包可能な一般状態空間モデルについても近年研究が進み、逐次モンテカルロ法の一つである粒子フィルタのアルゴリズムが与えられている。一般状態空間モデルと粒子フィルタに関しては、Doucet et al. (2001), 北川 (2005), 樋口 (2011)などに詳しく解説されている。

2.3 潜在変数を含む研究

マーケティングでは、POSデータやID付POSデータ、WEBの行動履歴等大規模かつ多次元のデータが蓄積されるようになった。これらビッグデータは大規模であったとしても消費者の行動の一端を測定しているに過ぎないといった認識をもたなければならない。消費者の行動は観測されている変数のみでは説明できず、その背後に存在する潜在的な変数や潜在的な構造までをも含めてモデル化、評価しなければならないのである。

Terui and Dahana (2006) (参照価格, 価格閾値), Terui and Ban (2008) (広告ストック), Terui et al. (2011) (広告ストック, 考慮集合), 井上 (2010) (ディテールストック), 佐藤・樋口 (2008a) (参照価格), 佐藤・樋口 (2009) (家庭内在庫量, 消費量), 山田・佐藤 (2012) (パフォーマンス評価/期待), 宮津・佐藤 (2015) (心理的財布), 青柳・佐藤 (2015) (参照価格, 広告ストック)等

表1. 周辺研究の概観.

文献	枠組み	消費者異質性	時間的異質性	潜在変数
Allenby and Rossi (1998)	階層ベイズモデル	○		
青柳・佐藤 (2015)	状態空間モデル		○	○
Bass et al. (2010)	状態空間モデル		○	
Bruce (2008)	状態空間モデル	○	○	
井上 (2010)	階層ベイズモデル	○		○
日高・佐藤 (2016)	階層ベイズモデル	○		
宮津・佐藤 (2015)	階層ベイズモデル	○		○
本橋・樋口 (2013)	状態空間モデル		○	
本橋 他 (2012)	状態空間モデル		○	
Kondo and Kitagawa (2000)	状態空間モデル		○	
Manchanda, and Chintagunta (2004)	階層ベイズモデル	○		
Manchanda et al. (2004)	階層ベイズモデル	○		
Naik et al. (1998)	状態空間モデル		○	
Rossi and Allenby (2003)	階層ベイズモデル	○		
佐藤・樋口 (2008a)	状態空間モデル		○	○
佐藤・樋口 (2008b)	状態空間モデル	○	○	
佐藤・樋口 (2009)	状態空間モデル	○	○	○
Terui and Ban (2008)	階層ベイズモデル	○		○
Terui and Dahana (2006)	階層ベイズモデル	○		○
Terui et al. (2011)	階層ベイズモデル	○		○
山口 他 (2004)	状態空間モデル		○	
山田・佐藤 (2012)	階層ベイズモデル	○		○
山田・佐藤 (2016)	階層ベイズモデル	○		

の研究では、前段に示した問題意識のもと、消費者異質性あるいは時間的異質性の概念に加えて、()内に示すような潜在変数や潜在構造を取り込んだモデル化がなされている。個々の潜在変数に関してここでは説明を割愛するが、それら変数、構造を捉えることが出来ればマーケティング実務の戦略構築に有用な知見を提供できる。その意味で、本項に示した事項はマーケティング分野やサービス分野で深化させなければならない領域である。

2.4 先行研究のまとめ

表1には、2.1項～2.3項に示した研究を総括的に整理した。紹介した研究のみではあるが、消費者異質性、時間的異質性、潜在変数のいずれか、あるいは複数が研究の視点として含まれている。これらの研究群は、マーケティングを対象としたものであるが、サービス研究に対しても重要な示唆を供給する。重要な点なので注記しておく。

3. 事例紹介

本節では、ベイジアンモデリングにより消費者の深い理解を狙いとして実施した既存研究(宮津・佐藤, 2015; 青柳・佐藤, 2015)のエッセンスを紹介する。

3.1 宮津・佐藤 (2015)

3.1.1 研究の概要

当該研究は、心理的財布と関連した消費者の心的状況を考慮した購買点数(バスケットサイズ)の生起メカニズムをモデル化し、その現象を明らかにすることを目的として実施した。当該研究では、2節で議論した「消費者異質性」を個人ごとのパラメータとして、さらに心的負荷

を「潜在変数」として表現したモデル化となっている。具体的には①消費者の購買時の心的状況を心的負荷として表現し、②その逼迫状況を心的負荷と閾値パラメータの大小関係として表現する階層ベイズ閾値ポアソン回帰モデルを提案している。

3.1.2 購買点数のモデル

消費者の小売店での購買点数(1度の買い物でバスケットに入る商品点数)は、小売店頭での値引きなどのプロモーション活動や慣性行動(平日に来店しやすい、週末に来店しやすい等)に影響されて規定される。ID付POSデータを用い、観測変数だけで消費者異質性に配慮しなければ容易にモデル表現、推定できる。しかし、消費者の買い物行動では、「今月はちょっと使いすぎたから出費を抑えよう」や「今日は給料日だからちょっとだけ贅沢しよう」といった心理的な状況の変化で、購買量や購買の質に変化が生じる。しかし、そういった心理的な状況は観測データで直接測定できず、何らかの形でモデル表現しなければ、必要な情報抽出はできず、さらに消費者行動を妥当に評価できない。本研究は、そのモチベーションのもとで実施したものである。以降では、本研究のモデルを中心にキーとなる事項だけに限定し、説明する。

心的負荷のモデル はじめに本研究の肝である消費者の心的負荷のモデル化を説明する。消費者個々は、前述したように、物理的な財布のほかにも心理的な財布を持っていると考えられている(Thaler, 1985)。しかし、それは単純に計量化できるものではない。本研究では、給料日からの累積購買金額を基に心的負荷のモデル化を行った。ただし、このアイデアには一つの課題がある。いつが給料日なのか?に関して、データが存在しないのである(そもそも顧客ごとにそれを調査することもできない)。その課題に対応するために、本研究では(3.1)式に示すように代表的な3つの給料日(25日, 5日, 17日)を設定する。

$$(3.1) \quad \text{累積購買集計期間} \begin{cases} l=1, \text{ 前月 25 日から今月 24 日まで} \\ l=2, \text{ 前月 5 日から今月 4 日まで} \\ l=3, \text{ 前月 17 日から今月 16 日まで} \end{cases}$$

(3.2)式は、消費者*i*の累積購買集計期間*l*における購買期間*t_i*までの累積購買金額を示す。

$$(3.2) \quad cumm_{i,t_i,l} = \begin{cases} \sum_{j=1}^{trans^l(t_i)} M_{i,j}, & trans^l(t_i) \neq 1 \\ 0, & trans^l(t_i) = 1 \end{cases}$$

$trans^l(t_i)$ は、消費者*i*の集計期間*l*における集計起点日から*t_i*までの来店回数を示す。 $M_{i,j}$ は消費者*i*の購買機会*j*における購買金額を示す。本研究では心的負荷を(3.3)式で表現する。

$$(3.3) \quad CummM_{i,t_i} = \alpha_i^{*(1)} cumm_{i,t_i,1} + \alpha_i^{*(2)} cumm_{i,t_i,2} + \alpha_i^{*(3)} cumm_{i,t_i,3}$$

$\alpha_i^{*(k)}, k=1,2,3$ は $0 \leq \alpha_i^{*(k)} \leq 1$ および $\sum_{k=1}^3 \alpha_i^{*(k)} = 1$ の制約を満たすパラメータであり、 $\alpha_i^* = (\alpha_i^{*(1)}, \alpha_i^{*(2)})$ とする(2つのパラメータが決まれば制約から3つ目のパラメータは自動的に決まる)。この定式化により、例えば $\alpha_i^{*(1)} = 1$ と推定された場合、消費者*i*の世帯は給料日が25日であると、 $\alpha_i^{*(1)} = 0.5, \alpha_i^{*(2)} = 0.5$ と推定された場合、消費者*i*の世帯には異なる給料日(25日と5日)の人がいる(ダブルインカム)等々、類推できることになる。

個体内モデル(観測モデル) (3.4)式が、本研究における個体内モデル(観測モデル)である。 y_{i,t_i} は消費者*i*の購買機会*t_i*の購買点数を示す。また、 $z_{i,t_i}^{(k)}, \beta_i^{(k)}$ は、消費者*i*の購買機会*t_i*の第*k*レジームの説明変数ベクトルと消費者*i*の第*k*レジームの反応係数を示す(回帰構造に関しては(3.5)式)。 y_{i,t_i} は非負のカウントデータであるため、(3.4)式に示すように観測モデル

の基本構造にはポアソン回帰モデルを採用する．ただし，モデルには閾値パラメータ γ_i を導入し，(3.3)式で定式化した心的負荷と比較することで，連続的にモデルが切換る構造を表現する． $CummM_{i,t_i} \geq \gamma_i$ のときは心的負荷が閾値を超えていることを示し，心的負荷が高い状態にあること(レジーム 1)を，一方で $CummM_{i,t_i} < \gamma_i$ のときは心的負荷が閾値を下回っていることを示し，心的負荷が低い状態にあること(レジーム 2)を表現する．すなわち，このモデル表現により心的負荷の状況により購買点数の生起メカニズムが切換る構造を定式化できたことになる．

$$(3.4) \quad \Pr(Y_{i,t_i} = y_{i,t_i} | \lambda_{i,t_i}^{(1)}, \lambda_{i,t_i}^{(2)}, \gamma_i, \alpha_i^*) = \begin{cases} \frac{(\lambda_{i,t_i}^{(1)})^{y_{i,t_i}} \exp(-\lambda_{i,t_i}^{(1)})}{y_{i,t_i}!}, & CummM_{i,t_i} \geq \gamma_i \\ \text{Regime1: 心的負荷が高い状態} \\ \frac{(\lambda_{i,t_i}^{(2)})^{y_{i,t_i}} \exp(-\lambda_{i,t_i}^{(2)})}{y_{i,t_i}!}, & CummM_{i,t_i} < \gamma_i \\ \text{Regime2: 心的負荷が低い状態} \end{cases}$$

$$(3.5) \quad \log(\lambda_{i,t_i}^{(k)}) = z_{i,t_i}^{(k)t} \beta_i^{(k)}, k = 1, 2$$

以降には得られた重要な知見を概説する．なお，モデルの全体，推定法の詳細は宮津・佐藤(2015)を確認してほしい．

3.1.3 得られた知見

本提案モデルの推定結果からは，様々な議論が出来る．しかしながら，紙幅の都合もあり，本研究のポイントでもある $\alpha_i^{*(k)}$, $k = 1, 2, 3$ と γ_i の推定結果のエッセンスのみを紹介する．

表 2 には，消費者ごとに算出した閾値パラメータ (γ_i) と心的負荷の構成パラメータ (α^*) の事後平均(サンプルサイズ $N = 1,000$)の基本統計量を示した．閾値パラメータの事後平均の平均は 1.7 万円であり，消費者の平均累積購買金額が 2.5 万円/月であることを考慮すれば，平均累積購買金額の 68% を超えると心理的財布の切換が生じる換算となる．また，構成パラメータの各平均値は 0.35 程度であり，各累積購買金額からの心的負荷への寄与は全体的に見ればほぼ均等である．図 2 には， $\alpha^{*(1)}, \alpha^{*(2)}$ の事後平均の分布状況を示した．図中 (1, 0), (0, 1), (0, 0) 近傍のプロットは購買期間がそれぞれ 1, 2, 3 のみ， $\alpha^{*(1)} + \alpha^{*(2)} = 1$ の直線上のプロットは購買期間が 1 と 2， $\alpha^{*(1)}$ 軸上のプロットは購買期間が 2 と 3， $\alpha^{*(2)}$ 軸上のプロットは購買期間が 1 と 3，図中の三角形内のプロットは全ての購買期間で構成されるものであり，消費者ごとにその構成に差が生じている．

表 3 には，構成パラメータの推定値から，心的負荷の構成パターンの割合を示した．ただし，各購買期間で 1% に満たないと推定されたものは切り捨てて算出している．ここで各構成要素は，心的負荷を構成している期間における累積購買金額の重みを表している．本モデルの設定

表 2. 推定値の基本統計量 ($N = 1,000$).

統計量	γ	$\alpha^{*(1)}$	$\alpha^{*(2)}$
平均値	0.1713	0.3538	0.3605
中央値	0.1073	0.0087	0.0079
最大値	1.8238	1.0000	1.0000
最小値	0.0055	0.0000	0.0000
標準偏差	0.2014	0.4347	0.4376

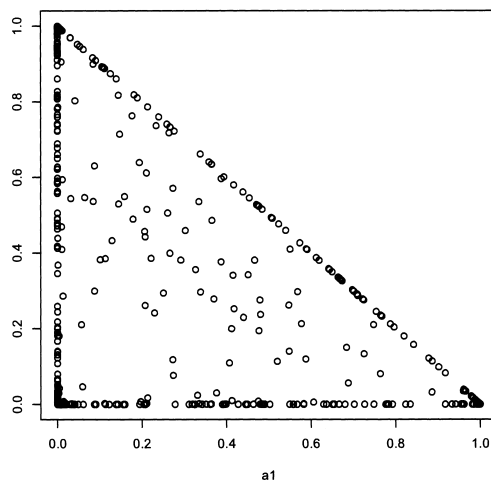
図 2. 構造パラメータ ($\alpha^{*(1)}, \alpha^{*(2)}$) の分布.

表 3. 構造パラメータの割合.

心理的財布の構成	割合 (%)	($\alpha^{*(1)} : \alpha^{*(2)} : \alpha^{*(3)}$)
購買期間 1 のみ	24.8	1.00 : 0.00 : 0.00
購買期間 2 のみ	24.6	0.00 : 1.00 : 0.00
購買期間 3 のみ	16.2	0.00 : 0.00 : 1.00
購買期間 1 と 2	8.4	0.58 : 0.42 : 0.00
購買期間 2 と 3	9.4	0.00 : 0.51 : 0.49
購買期間 3 と 1	9.3	0.52 : 0.00 : 0.48
全ての購買期間	7.3	0.32 : 0.36 : 0.32

では、期間 1, 2, 3 は、給料日がそれぞれ 25 日, 5 日, 17 日に相当する。異なる給与支給日からの累積購買金額で心的負荷が構成される場合、世帯に複数の給与と所得者が存在すると推察できる。実証分析に基づけば、心的負荷が各期間単独で構成される世帯が全体の 65% を占めており、単独ないしは複数の給与と所得者が存在しても同一給料日である世帯がこれに相当すると考えられる。35% 近くは心的負荷の構成に複数の累積購買金額が影響している。これらは、例えばパートタイムを含む給料日が異なる共働き世帯の構造を反映している。

本研究は、心的負荷をモデル化し、観測モデルに閾値構造を取り込むことにより、観測データを単純に用いるだけでは獲得しえない知見を抽出できた。消費者を理解するためには、こういった統計モデルの技術が有用であるとイメージしてもらえればと考える。

3.2 青柳・佐藤 (2015)

3.2.1 研究の概要

当該研究は、セールスプロモーションの売上に対する動的効果をテレビ広告ストックと参照価格で階層化し、その動的進展メカニズムの解明を目的として実施した。当該研究では、2 節で議論した「時間的異質性」を時変パラメータとして、さらに集計レベルの参照価格とテレビ広告ストックを「潜在変数」として表現したモデル化となっている。具体的には、①集計レベルの

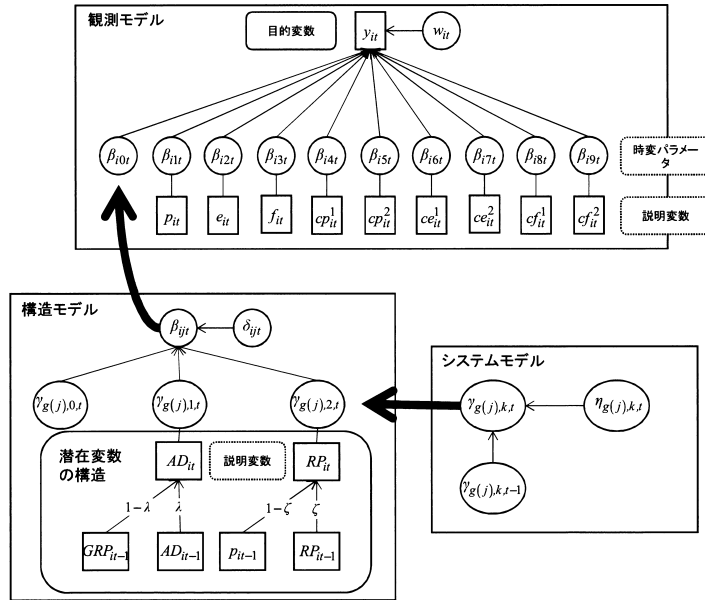


図 3. モデルの全体像.

対数点数PI(来店客千人当たり販売点数：3商品)をセールスプロモーション(価格, 山積み, チラシ：自身, 競合)で説明する動的多変量回帰(観測モデル), ②①の時変係数を目的変数とし, 広告ストックおよび参照価格(両変数とも潜在変数)で説明する動的多変量回帰(構造モデル), ③②の動的多変量回帰の時変係数の時間進展メカニズム(システムモデル), の3つをモデル化している.

3.2.2 動的集計型市場反応モデル

モデルの全体像 本研究におけるモデルは, 外形的に非常に複雑に見えるため, はじめに図によりモデルの全体像を説明する. 図3にはモデルの全体像を示した. 本研究のモデルは, 基本構造として通常の状態空間モデル(観測モデル+システムモデル)を3層の状態空間モデル(観測モデル+構造モデル+システムモデル)に拡張したものである. 3階層状態空間モデルにおいて, 観測モデルとシステムモデルの役割は, 基本的に通常の状態空間モデルと同様である. 構造モデルは, 観測モデルに含まれる時変係数の時間進展の「理由」を記述するモデルであり, このモデル化が本研究のポイントになっている. マーケティング分野において時変係数をモデル化する場合, 平滑化事前分布を用いることが多い. そのアプローチは, 簡便に時変係数を表現できる一方で, 時間変化のメカニズムが分からないといった批判にさらされることが多い. 本提案モデルはその批判に応えるアプローチとなっている. 以降には, 提案モデルの個々をごく簡単に説明する. なお, 本研究は3商品($i = 1, 2, 3$)を対象としている. 表4には観測モデルで用いた変数を示す.

観測モデル 観測モデルは, 動的市場反応を多変量回帰モデルの枠組みで表現する. (3.6)式は商品 i ($i = 1, 2, 3$) 個々の動的市場反応モデルを示す. 実際には(3.6)式の3商品をベクトル表現し, 解析に用いる.

表 4. 観測モデルの変数一覧.

記号	内容
y_{it}	点数 PI (対数)
p_{it}	自商品価格掛率 (対数)
e_{it}	自商品エンド陳列実施
f_{it}	自商品チラシ掲載
cp_{it}^1	競合商品 1 価格掛率 (対数)
cp_{it}^2	競合商品 2 価格掛率 (対数)
ce_{it}^1	競合商品 1 エンド陳列実施 (対数)
ce_{it}^2	競合商品 2 エンド陳列実施 (対数)
cf_{it}^1	競合商品 1 チラシ掲載
cf_{it}^2	競合商品 2 チラシ掲載
β_{ijt}	観測モデルにおける回帰係数 (市場反応係数)
w_{it}	観測ノイズ
σ_i^2	観測ノイズの分散

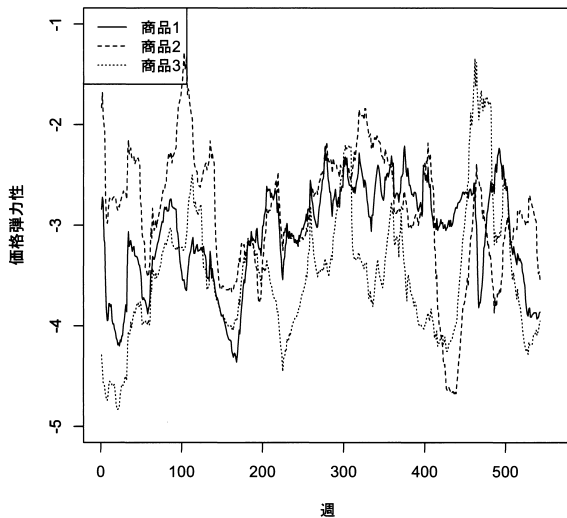


図 4. 動的価格弾力性.

$$(3.6) \quad y_{it} = \beta_{i0t} + p_{it}\beta_{i1t} + e_{it}\beta_{i2t} + f_{it}\beta_{i3t} + cp_{it}^1\beta_{i4t} + cp_{it}^2\beta_{i5t} + ce_{it}^1\beta_{i6t} + ce_{it}^2\beta_{i7t} + cf_{it}^1\beta_{i8t} + cf_{it}^2\beta_{i9t} + w_{it}, w_{it} \sim N(0, \sigma_i^2)$$

構造モデル 構造モデルは、図 4 に持式的に示したように β_{ijt} が自身の広告ストックと参照価格 AD_{it}^{own} , RP_{it}^{own} および競合の広告ストックおよび参照価格, AD_{it}^{comp} , RP_{it}^{comp} それぞれから影響を受け、時間発展する様子を表現する。(3.7)式が提案の構造モデルになる。

$$(3.7) \quad \beta_{ijt} = \log(AD_{it}^{own})\gamma_{g(j),1,t} + \log(RP_{it}^{own})\gamma_{g(j),2,t} + \log(AD_{it}^{comp})\gamma_{g(j),3,t} + \log(RP_{it}^{comp})\gamma_{g(j),4,t} + \delta_{ijt}, \delta_{ijt} \sim N(0, \tau_{ij}^2)$$

(3.7)式で用いる広告ストックと参照価格は、自商品分、競合商品分を区別して AD_{it}^{own} , AD_{it}^{comp} と添え字を用いて表現している。以降の説明では簡単のために、これらを区別せず AD_{it} , RP_{it} と記載する。(3.8)式が広告ストックの、(3.9)式が参照価格の更新式を示す。

$$(3.8) \quad AD_{it} = \lambda AD_{i,t-1} + (1 - \lambda) GRP_{i,t-1}$$

$$(3.9) \quad RP_{it} = \zeta RP_{i,t-1} + (1 - \zeta) Price_{i,t-1}$$

λ , ζ は、更新の程度を規定する平滑化パラメータで、0 以上 1 以下の値をとるものと仮定する。この値が 1 に近いほど前の時点のストックが残存し、逆に 0 に近いほど残存しないことを意味する。

システムモデル (3.10)式は、変数 k の時間進展を示すシステムモデルである。 $g(j)$ が同じ数値になっている場合、(3.10)式は共通のものを用いる。本研究では、システムモデルを平滑化事前分布の考え方に基づき、滑らかさの仮定のもとでモデル化する。

$$(3.10) \quad \gamma_{g(j),k,t} = \gamma_{g(j),k,t-1} + \eta_{g(j),k,t}, \eta_{g(j),k,t} \sim N(0, \xi_{g(j),k}^2)$$

3 階層状態空間モデル 本研究では、上述の観測モデル、構造モデル、システムモデルは 3 階層状態空間モデルの枠組みで統合できる。3 階層状態空間モデルは(3.11)式～(3.13)式の 3 つのモデルで表現されるモデルであり、(3.6)式～(3.10)式のモデルをベクトル表現すれば得られる。その詳細については、青柳・佐藤 (2015)を参照してほしい。

$$(3.11) \quad \mathbf{y}_t = \mathbf{H}_{1,t} \mathbf{x}_{1,t} + \mathbf{w}_{1,t}, \mathbf{w}_{1,t} \sim \text{MVN}(\mathbf{0}, R_1), \quad (\text{観測モデル})$$

$$(3.12) \quad \mathbf{x}_{1,t} = \mathbf{H}_{2,t} \mathbf{x}_{2,t} + \mathbf{w}_{2,t}, \mathbf{w}_{2,t} \sim \text{MVN}(\mathbf{0}, R_2), \quad (\text{構造モデル})$$

$$(3.13) \quad \mathbf{x}_{2,t} = \mathbf{x}_{2,t-1} + \mathbf{w}_{3,t}, \mathbf{w}_{3,t} \sim \text{MVN}(\mathbf{0}, R_3), \quad (\text{システムモデル})$$

3.2.3 得られた知見

本提案モデルの推定結果からは、様々な議論が出来る。本研究のポイントでもある商品ごとの動的価格弾力性 β_{i1t} , $i = 1, 2, 3$ とその動的構造パラメータ $\gamma_{2,k,t}$, $k = 1, \dots, 4$ の推定結果に絞って、そのエッセンスを紹介する。

本研究では、インスタントカレーの 3 商品のデータを用いた。分析には 3 商品のもので POS データとコーザルデータ(売価、山積み陳列、チラシ掲載、広告投下量(GRP))を用いた。表 5 には、変数ごとの要約統計量を示した。

表 5. 使用データの要約統計量.

項目	商品 A	商品 B	商品 C
平均販売個数	3.357	8.101	5.085
平均点数 PI	1.182	1.772	2.449
販売個数欠測日数	109	106	33
最大売価(定価)	218	198	198
平均価格掛率	0.867	0.824	0.855
エンド陳列回数	72	161	119
チラシ掲載回数	7	19	28
GRP 総和	4,809	4,915	19,457

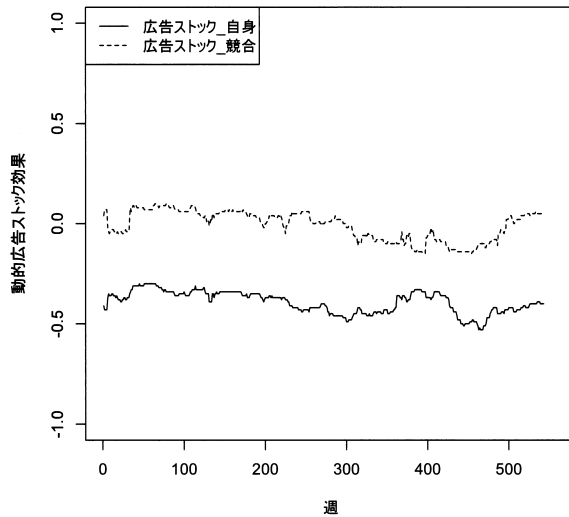


図 5. 動的広告ストック効果.

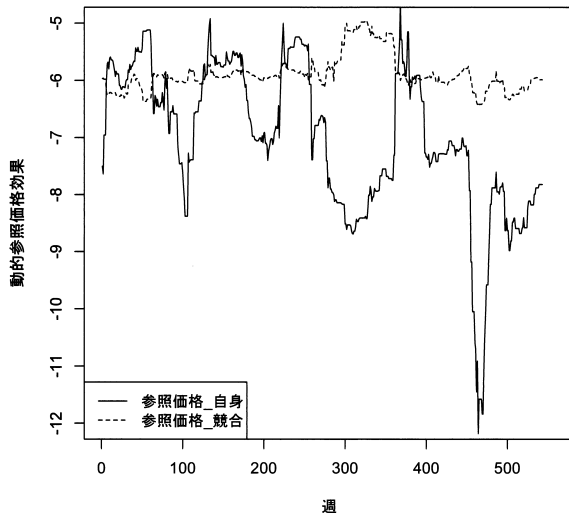


図 6. 動的参照価格効果.

図 3 には、推定した商品ごとの価格弾力性消費者 $\beta_{i1t}, i = 1, 2, 3$ を示した。商品ごとにその推移に差が生じている。時間的異質性を捉えるという観点で、時変価格弾力性は十分に意味がある。しかし、マーケティングを高度化するにはさらなる踏み込みが必要である。上述しているが、 $\beta_{i1t}, i = 1, 2, 3$ の変動が生じる理由をも評価したいのである。そこで重要なのが、(3.7) 式に示した構造モデルである。図 5 には広告ストック(自身、競合)の $\beta_{i1t}, i = 1, 2, 3$ に与える動的影響を、図 6 には参照価格(自身、競合)の $\beta_{i1t}, i = 1, 2, 3$ に与える動的影響をそれぞれ示した。これら図から、自商品の広告ストックが増えるほど、また自商品の店舗レベル参照価格が低下するほど、価格弾力性が負に大きくなる。ここに示したような、時間的異質性を捉え、

しかもその変動理由をも同時に捉えることは、実務マーケティングにおける「何を、何のために、コントロールするのか」といった課題に対して示唆を与えられる。今回の事例でいえば、その構造は、「価格や広告投下量は参照価格や広告ストックの変動に影響し、それら潜在変数がセールスプロモーションの動的市場反応に影響する。最終的にはその市場反応により売上が規定される」というものに対応する。これは通常考えられる「売価や広告が直接売上が規定する」といったものと大きく異なる。消費者の行動を理解し、それに応じた形式でマーケティングを高度化するには、理由までも含めたモデリングが必要だし、有効だと考える。本研究は一つの事例でしかないが、消費者の理解を深めるには、この種のアプローチを進展させなければならない。

4. まとめと今後の課題

本研究は、消費者理解を深めるにはという観点で統計モデル(特にベイズモデル)をどのように活用すべきか?に対する提言を目的とし、既存研究の紹介を基本として整理した。現象が変われば必要なモデリングに違いが生じるが、本稿の内容は統計サイドから研究対象が主として消費者のサービスやマーケティング分野に接近する際の参考にしてもらえるのではないかと考えている。

サービス研究の源流は、ほぼマーケティング研究にあるものと考えられる。マーケティング研究から派生したサービス研究は、その重要性が叫ばれ、日本学術会議におけるサービスの参照規準小委員会でも「人間がその系に含まれる連続的システムにおいて、感情や知識を含む様々な価値を共創的かつダイナミックに生産する行為」といった形式でサービスの定義を与えている。伝統的なマーケティングが表6に示す Goods Dominant Logic(GDL)の考え方に基づいている一方で、サービスは表6に示す Service Dominant Logic(SDL)の考えに基づき進展している。その進展における重要なポイントは、GDLが一方方向の価値提供なのに対し、SDLは双方方向の価値共創になっている点である。SDLの考え方に基づくと価値は供給者サイドが生み出すものではなく、供給者と消費者が共創することによって生み出されるのだと、考えることになる。この発想自体が誤りであるわけではないが、共創という発想を強く意識するがゆえに、本来供給者サイドがしなければいけない「消費者の理解」がなおざりになってしまっている。価値共創の仮定のもとでも供給者は消費者理解を進めなければならないし、それをしない限り競争優位性は構築できない。その際有効に機能するのが本項で紹介したアプローチであり、統計モデルである。サービス科学分野では、本稿で指摘した事項を意識した形で消費者理解のための研究を進めなければならない。また、それが実現できれば社会に大きな貢献をもたらすものと確信している。

表6. Goods Dominant Logic と Service Dominant Logic.

視点	GDL : Goods Dominant Logic 一方方向な価値提供 「交換価値」の最大化	SDL : Service Dominant Logic 双方方向の価値共創 「使用価値」の最大化
価値の生産	企業	企業と顧客が共に
顧客の役割	価値を消費する	価値を生産し、消費する
イノベーションの対象	製品や技術	顧客とのインタラクション
価値の源泉	新しい機能や性能(交換価値)	新しい顧客行動や顧客経験(使用価値)
企業と顧客との接点	購買時	購買前, 購買時, 購買後

参 考 文 献

- Allenby, G. M. and Rossi, P. E. (1998). Marketing models of consumer heterogeneity, *Journal of Econometrics*, **89**(1), 57–78.
- 青柳憲治, 佐藤忠彦 (2015). 3階層多変量状態空間モデルによる動的市場反応形成メカニズムの解明, 日本オペレーションズ・リサーチ学会和文論文誌, **58**, 70–100.
- Bass, F., Bruce, N., Majumdar, S. and Murthi, B. (2007). Wearout effects of different advertising themes: A dynamic Bayesian model of the advertising-sales relationship, *Marketing Science*, **26**(2), 179–195.
- Bruce, N. (2008). Pooling and dynamic forgetting effects in multitheme advertising: Tracking the advertising sales relationship with particle filters, *Marketing Science*, **27**(4), 659–673.
- Chib, S. and Greenberg, E. (1996). Markov chain Monte Carlo simulation methods in econometrics, *Econometric Theory*, **12**(3), 409–431.
- Doucet, A., Freitas, N. de and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*, Springer, New York (etc.).
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**(1), 97–109.
- 日高徹司, 佐藤忠彦 (2016). 消費者とブランドとの関係を考慮した階層ベイズモデルによるクロスメディア効果推定, 日本オペレーションズ・リサーチ学会和文論文誌, **59**, 106–133.
- 樋口知之 (2011). 『予測にいかす統計モデリングの基本—ベイズ統計入門から応用まで』, 講談社, 東京.
- 井上友彦 (2010). 医師の異質性を考慮した医薬品業界における営業訪問効果の分析, マーケティングサイエンス, **18**, 49–73.
- 北川源二郎 (2005). 『時系列解析入門』, 岩波書店, 東京.
- Kondo, F. N. and Kitagawa, G. (2000). Time series analysis of daily scanner sales: Extraction of trend, day-of-the-week effect and price promotion effect, *Marketing Intelligence & Planning*, **18**(2), 53–66.
- Manchanda, P. and Chintagunta, P. (2004). Responsiveness of physician prescription behavior to sales-force effort: An individual level analysis, *Marketing Letters*, **15**(2), 129–145.
- Manchanda, P., Rossi, P. and Chintagunta, P. (2004). Response modeling with nonrandom marketing-mix variables, *Journal of Marketing Research*, **41**(4), 467–478.
- 宮津和弘, 佐藤忠彦 (2015). 心理的財布と購買行動の関係性のモデル化—階層ベイズ閾値ポアソン回帰モデルの提案, 応用統計学, **44**(3), 161–182.
- 本橋永至, 樋口知之 (2013). 市場構造の変化を考慮したブランド選択モデルによる購買履歴データの解析, マーケティング・サイエンス, **21**(1), 37–59.
- 本橋永至, 磯崎直樹, 長尾大道, 樋口知之 (2012). 状態空間モデルによるインターネット広告のクリック率予測, オペレーションズ・リサーチ: 経営の科学, **57**(10), 574–583.
- Naik, P., Mantrala, M. and Sawyer, A. (1998). Planning media schedules in the presence of dynamic advertising quality, *Marketing Science*, **17**(3), 214–235.
- 日本学術会議経営学委員会・総合工学委員会合同サービス学分会 (2017). 報告, 大学教育の分野別質保証のための教育課程編成上の参照基準, サービス学分野, 1–28.
- Rossi, P. and Allenby, G. (2003). Bayesian statistics and marketing, *Marketing Science*, **22**(3), 304–328.
- Rossi, P., Allenby, G. and McCulloch, R. (2005). *Bayesian Statistics and Marketing*, John Wiley & Sons., England.
- 佐藤忠彦, 樋口知之 (2008a). 動的売上反応モデルによる POS データの解析, マーケティングサイエンス, **15**(1), 1–26.
- 佐藤忠彦, 樋口知之 (2008b). 動的個人モデルによる消費者来店行動の解析(議論付き), 日本統計学会論文誌, **38**(1), 1–38.
- 佐藤忠彦, 樋口知之 (2009). 動的個人モデルによる購買生起行動の解析, マーケティング・サイエンス,

- 16(1, 2), 49–73.
- 佐藤忠彦, 樋口知之 (2013). 『ビッグデータ時代のマーケティング—ベイジアンモデリングの活用』, 講談社, 東京.
- 照井伸彦 (2008). 『ベイズモデリングによるマーケティング分析』, 東京電機大学出版局, 東京.
- Terui, N. and Ban, M. (2008). Modeling heterogeneous effective advertising stock using single-source data, *Quantitative Marketing and Economics*, **6**(4), 415–438.
- Terui, N. and Dahana, W. D. (2006). Research note—Estimating heterogeneous price thresholds, *Marketing Science*, **25**(4), 384–391.
- Terui, N., Ban, M. and Allenby, G. M. (2011). The effect of media advertising on brand consideration and choice, *Marketing Science*, **30**(1), 74–91.
- Thaler, R. (1985). Mental accounting and consumer choice, *Marketing Science*, **4**, 199–214.
- 山口類, 土屋映子, 樋口知之 (2004). 状態空間モデルを用いた飲食店売上の要因分解, オペレーションズ・リサーチ: 経営の科学, **49**(5), 52–60.
- 山田浩喜, 佐藤忠彦 (2012). 階層ベイズモデルによる百貨店の態度ベース店舗満足化構造に関する解析, マーケティングサイエンス, **20**(1), 17–41.
- 山田浩喜, 佐藤忠彦 (2016). 百貨店顧客の来店回数生起メカニズムの構造異質性の解析, 行動計量学, **43**(1), 53–68.

Possibility of Achieving Consumer Understanding Using Statistical Models

Tadahiko Sato

Faculty of Business Sciences, University of Tsukuba

Advanced utilization of statistical models is an indispensable tool for the development of service sciences, and its importance will increase further in the future. The purpose of this paper is to explain how to use a statistical model (specifically, a Bayesian model) to achieve deep consumer understanding, which is essential for improving services. Specifically, we will organize related research on this issue and provide details of our two research. Although these two studies were in marketing research, they also provide suggestions relevant to service research.