

データ同化による消費者の購買行動分析に向けて

総合研究大学院大学 統計科学専攻
博士課程 本橋永至

1 はじめに

近年、企業は POS データ、スキャナ・パネル・データ、広告シングル・ソース・データ、FSP (Frequent Shopper Program) による会員顧客の購買履歴データなど大量のマーケティング・データを容易に収集できるようになった。また、情報処理技術の進歩により個人レベルのデータを平均や分散などに集約してからデータを分析するのではなく、非集計のデータをそのまま分析して個々の顧客の行動を把握することが可能になってきた。一方、消費者の嗜好の多様化や競合企業との競争の激化などの理由から、新規顧客の獲得よりも既存顧客の育成を重視する CRM (Customer Relationship Management) やワン・トゥ・ワン・マーケティングといったアプローチが広まりつつある。本研究では、消費者の購買時における考慮集合の変化のメカニズムを状態空間アプローチにより表現したブランド選択モデルを提案する。

2 モデル

2.1 ブランド選択モデル

ブランド選択モデルの構築に際して以下の仮定を設ける。あるスーパー・マーケットの消費者は総計 K 人おり、各消費者は価格と特別陳列の有無を考慮して、ある商品カテゴリーの購買ブランドを選択する。ただし、消費者は当該カテゴリーのすべてのブランド I 個を選択の対象とするのではなく、各消費者が購買機会毎に持つ考慮集合に入っているブランドのみを選択の対象とする。

ブランド i が第 t 日において消費者 k の考慮集合に入っている確率を

$$(2.1) \quad S_{ik}^t = \frac{\exp(\alpha_{ik}^t)}{1 + \exp(\alpha_{ik}^t)}$$

とする。 α_{ik}^t は、第 t 日におけるブランド i の消費者 k の考慮集合への入りやすさを表す状態変数である。従って、消費者 k の第 t 日における考慮集合が C である確率は、

$$(2.2) \quad P_k^t(C) = \prod_{i \in C} S_{ik}^t \prod_{i \notin C} (1 - S_{ik}^t)$$

と表せる。さらに、消費者 k の第 t 日における考慮集合 C からのブランド i の選択確率を多項ロジットモデルによって

$$(2.3) \quad P_k^t(i|C) = \frac{\exp V_{ik}^t}{\sum_{j \in C} \exp(V_{jk}^t)}$$

のように定式化する (McFadden 1973)。 V_{ik}^t は消費者 k の第 t 日におけるブランド i の確定効用であり、

$$(2.4) \quad V_{ik}^t = \beta_{0i}^t + \beta_{1i}^t PRICE_i^t + \beta_{2i}^t FEATURE_i^t$$

のように線形の効用関数を仮定する。ここで、

$PRICE_i^t$ = 第 t 日のブランド i の販売価格,
 $FEATURE_i^t$ = 第 t 日のブランド i の特別陳列の有無

である。ブランド選択確率は、ある考慮集合 C の得られる確率とブランド i の C からの選択確率の積和として表現できるため、消費者 k の第 t 日におけるブランド i の選択確率を

$$(2.5) \quad P_k^t(i) = \sum_C P_k^t(i|C)P_k^t(C)$$

と定式化する。

2.2 状態空間表現

一般状態空間モデルは、観測モデルとシステムモデルの2つのモデルから定式化され、本研究のモデルは下記のように表現できる。

$$(2.6) \quad \text{観測モデル} \quad P_k^t(i) = \sum_C P_k^t(i|C)P_k^t(C)$$

$$(2.7) \quad \text{システムモデル} \quad x^t = Fx^{t-1} + Gv^t$$

ただし、 x^t 及び v^t は状態ベクトル、システム・ノイズ・ベクトルを示し、

$$(2.8) \quad x^t = (\alpha_{11}^t, \dots, \alpha_{IK}^t, \beta_{01}, \dots, \beta_{0I}, \beta_{11}, \dots, \beta_{1I}, \beta_{21}, \dots, \beta_{2I})$$

である。さらに、それぞれの状態変数の変化を、

$$(2.9) \quad \alpha_{ik}^t = \alpha_{ik}^{t-1} + \gamma_1 DAYSt_{ik}^t + \gamma_2 FEATURE_i^t + v_{\alpha_{ik}}^t$$

$$(2.10) \quad \beta_{0i}^t = \beta_{0i}^{t-1} + v_{\beta_{0i}}^t$$

$$(2.11) \quad \beta_{1i}^t = \beta_{1i}^{t-1} + v_{\beta_{1i}}^t$$

$$(2.12) \quad \beta_{2i}^t = \beta_{2i}^{t-1} + v_{\beta_{2i}}^t$$

と定式化する。ここで、 $DAYSt_{ik}^t$ は、消費者 k のブランド i の前回購買日からの日数である。

3 今後の方向性

スーパー・マーケットの ID 付き POS データを利用して本研究で提案したモデルの実証分析を行う。状態推定には粒子フィルタ/平滑化のアルゴリズムを用いる。本研究は消費者行動のメカニズムの解明、さらにはマーケティングにおける需要予測のための新しい分析手法を提示した。状態推定に粒子フィルタ/平滑化を用いた研究はマーケティング分野において始まったばかりであり(佐藤・樋口 2008)、これからその有効性がますます注目されることが予想される。今後は、クチコミや SNS などの消費者間の相互作用が消費者の購買行動に影響を及ぼすメカニズムをとらえるモデルの開発に取り組んでいきたい。

参考文献

- 佐藤忠彦, 樋口知之 (2008). 「動的個人モデルによる消費者来店行動の解析」, 『日本統計学会誌』 **38** (1), 1-19.
- McFadden, D. (1973). Conditional Logit Analysis in Qualitative Choice Behavior, in *Frontiers in Econometrics*, P. Zarembka, ed. Academic Press, New York.