ノンパラメトリック密度関数による マーケットシェアの推定

山上 伸, 北澤 英里子, 中村 肇

1. はじめに

消費者行動を予測するべく, 50年代よりさまざまな形のブランド選択モデルが提唱されてきた. 中でも, ロジットモデルは最も広く用いられている. 近年, 膨大な量のパネルデータがスキャナーから取り込まれ, 蓄積されるようになり, 消費者に関するデータが比較的安価に入手可能となってきた. そこで, 今回の解析では, オブザベーションが豊富であることを利用して, マーケティングミックスを変数とするノンパラメトリック手法による (Abe, 1991), 11種類のインスタントコーヒーのマーケットシェア予測を試みる.

この手法は, モデル構築の概念がわかりやすく, 自由度が大きいため, 歪み, 偏りを少なく表現することができる. また, 従来のパラメトリックモデルで必要とされるような前提条件も少なくてすむ.

2. 基本概念

ブランド j が選択される確率を, パラメータを用いた統計モデルを使わずに, スキャンパネルデータから直接推定することを考える. まず, あるスキャンパネルデータのマーケティングミックス M_i で, ブランド j が購入されたという実績から, M_i の近傍においても同じような購買行動が行なわれるであろうという仮定をおく. 次に, M_i と M_i の近傍 M との類似度をカーネル関数を用いて定義し, 類似度の値から M の下でブランド j が買われるという条件つき購買確率 $P(Y_j|M)$ を推定する.

$$P(Y_j|M) = \frac{P(Y_j, M)}{P(M)} = \frac{\sum_{i=1}^{N} y_{ij} K\left(\frac{M - M_i}{h}\right)}{\sum_{i=1}^{N} K\left(\frac{M - M_i}{h}\right)} \tag{1}$$

ここで,

$$y_{ij} = \begin{cases} 1, & i\text{がブランド}\,j\text{を購入した場合} \\ 0, & \text{その他の場合} \end{cases}$$

N :オブザベーションの数

h :平滑定数

$K(\cdot)$:カーネル関数

を表す.

カーネル関数については, モデルの性質上,

(a) 連続, 微分可能

(b) 積分すると1に等しい

という, 通常のノンパラメトリック推定で用いられる制約条件の他に,

(c) 非負, 単調減少

を制約条件に加えて, 適当な関数形を選ぶことにする.

3. ノンパラメトリック密度関数

前節のモデルを用い, マーケティングミックスの変数 M と, カーネル関数 $K(\cdot)$ の形を適当に選び, (1) によって推定される購入確率が最も当てはまる密度関数の構築を行なう.

始めにマーケティングミックスの変数として, 価格掛け率, エンド陳列の有無, チラシ宣伝の有無, 及びブランド選好における個人差を表すロイヤルティを用いる. ロイヤルティは, Guadagni & Little (1983) によって定義されたものと個人別のマーケットシェアを

やまがみ しん, きたざわ えりこ, なかむら はじめ 東京ガス(株)
〒260-91 千葉市美浜区中瀬 2-3

用いたものの2通りを考え，チューニングの結果，よりデータにフィットする方を採用する．

Guadagni & Little 方式：

$$L_{pj}(t) = \begin{cases} 1/J, & t=1\text{の場合} \\ \alpha \times L_{pj}(t-1) & \\ \quad +(1-\alpha)d_j(t-1), & t>1\text{の場合} \end{cases}$$

ここで，

$L_{pj}(t)$：パネル p のブランド j に対するロイヤルティ

$$d_j(t) = \begin{cases} 1, & \text{ブランド } j \text{ が選択される場合} \\ 0, & \text{ブランド } j \text{ が選択されない場合} \end{cases}$$

t：購買機会

J：ブランド数

α：衰退定数

を表す．この方式によれば，パネルが特定のブランドを買い続けた時にロイヤルティが上がる．また，衰退定数αは，ロイヤルティの変化の大きさを表す．

個人別マーケットシェア方式：

$$L_{pj} = \frac{\text{パネル } p \text{ のブランド } j \text{ の購入数}}{\text{パネル } p \text{ の全ブランドの購入数}}$$

次に，マーケティングミックス $M = (m_1, m_2, \ldots, m_d)$ と $M_i = (m_{i1}, m_{i2}, \ldots, m_{id})$ の類似度を表す $K(M-M_i, H)$ を，以下の式で与える．

$$K(M-M_i, H) = \prod_{l}^{d} k(m_l - m_{li}, h_l) \tag{2}$$

各マーケティングミックスの変数 m と m_iの類似度の $k(m-m_i, h)$ については，変数が連続値の場合を (3) で，バイナリーの場合を (4) で推定することにする．

$$k(m-m_i, h) = \frac{1}{\sqrt{2\pi h^2}} \exp\left\{-\frac{(m-m_i)^2}{2h^2}\right\} \tag{3}$$

$$k(m-m_i, h) = \mu\delta_{mm_i} + (1-\mu)(1-\delta_{mm_i}) \tag{4}$$

ここで，

h, μ：平滑定数

$$\delta_{xy} = \begin{cases} 1, & x=y\text{の場合} \\ 0, & x \neq y\text{の場合} \end{cases}$$

を表す．

今事例におけるカーネル密度関数の推定式を (5) に示す．

$$\begin{aligned} & K(M-M_i, H) \\ = & \prod_{j=1}^{11} \frac{1}{\sqrt{2\pi h_p^2}} \exp\left\{-\frac{(pri_j - pri_{j(i)})^2}{2h_p^2}\right\} \\ \times & \prod_{j=1}^{10} \frac{1}{\sqrt{2\pi h_l^2}} \exp\left\{-\frac{(loy_j - loy_{j(i)})^2}{2h_l^2}\right\} \\ \times & \prod_{j=1}^{11} \left[\mu\, \delta_{end_j\, end_{j(i)}} + (1-\mu)\left(1-\delta_{end_j\, end_{j(i)}}\right)\right] \\ \times & \prod_{j=1}^{11} \left[\mu\, \delta_{ad_j\, ad_{j(i)}} + (1-\mu)\left(1-\delta_{ad_j\, ad_{j(i)}}\right)\right] \end{aligned} \tag{5}$$

変数は，以下の通りである．

pri_j：ブランド j の価格

loy_j：ブランド j に対するロイヤルティ

ad_j：ブランド j に対するチラシの有無

end_j：ブランド j に対するエンド陳列の有無

h_p, h_l, μ：平滑定数

4. 平滑定数のチューニング

前節で与えたカーネル密度関数では，平滑定数をどのように設定するかという問題がまだ残されている．平滑定数は，類似度の近傍をどのくらいの大きさに取るのかを意味し，値が大きすぎる場合には推定精度が落ち，小さすぎる場合には頑健性が失われる．適当な平滑定数を得るために，以下のステップで，平滑定数のチューニングを行なう．チューニングに際して平滑定数は全ブランドで共通とし，エンド陳列とチラシの有無に対応する平滑定数も共通とする．

STEP1 平滑定数の適当な組合せを初期値として設定する．

STEP2 平滑定数とその近傍の組合せに対して，それぞれ 93 年 1 月から 9 月までのデータからモデルを作成する．

STEP3 93年10月から12月のシェアを予測し，タイルの不一致係数を求める．

STEP4 選択されたモデルが，近傍の中で不一致係数が最小な場合，これを採用する．

STEP5 そうでない場合，不一致係数が最小なモデルの平滑定数を選択して，STEP2 へ．

5. 結果

5.1 チューニング結果

Guadagni & Little 方式をロイヤルティとして採用したモデルの最終ステップにおける結果を表1に示す．

表 1: Guadagni & Little 方式

Case No.	h_p	h_l	μ	U
最良	0.05	0.05	0.90	0.069*
近傍1	0.10	0.10	0.95	0.078
近傍2	0.10	0.10	0.85	0.080
近傍3	0.10	0.03	0.95	0.151
近傍4	0.10	0.03	0.85	0.149
近傍5	0.03	0.10	0.95	0.070
近傍6	0.03	0.10	0.85	0.072
近傍7	0.03	0.03	0.95	0.070
近傍8	0.03	0.03	0.85	0.071

同様に，個人別マーケットシェア方式を採用したモデルのチューニング結果を表2に示す．

表 2: 個人別マーケットシェア方式

Case No.	h_p	h_l	μ	U
最良	0.05	0.20	0.90	0.051*
近傍1	0.10	0.25	0.95	0.066
近傍2	0.10	0.25	0.85	0.070
近傍3	0.10	0.10	0.95	0.090
近傍4	0.10	0.10	0.85	0.094
近傍5	0.03	0.25	0.95	0.052
近傍6	0.03	0.25	0.85	0.054
近傍7	0.03	0.10	0.95	0.062
近傍8	0.03	0.10	0.85	0.065

次に，タイルの不一致係数が最小である個人別マーケットシェアを採用し，最良の平滑定数の組合せの時の，実績値と推定値の比較を表3に示す．

表 3: 実績値と推定値の比較

ブランド番号	推定値%	実績値%
1	9.17	7.82
2	19.95	19.15
3	11.79	10.05
4	9.23	11.54
5	0.67	1.06
6	7.75	9.79
7	3.88	2.39
8	3.27	3.19
9	1.47	1.49
10	2.36	2.02
11	30.45	31.49

5.2 マーケットシェア推定結果

以上の結果に基づいて，93年度の1月から12月までのデータとチューニングされた平滑定数を用いて，モデルを構築し，94年度のマーケットシェアの予測を行なった．推定値と実測値を表4に示す．この結果，

タイルの不一致係数：0.071

という評価値が得られた．特に，メジャーブランドである1〜4，6〜9においては，相対誤差±10％以内で，シェアの予測に成功している．

表 4: マーケットシェアの予測結果

ブランド番号	推定値%	実績値%
1	16.73	14.90
2	12.43	11.48
3	8.80	10.05
4	11.61	9.21
5	0.89	0.21
6	9.09	8.58
7	4.17	4.33
8	2.00	2.55
9	4.34	5.37
10	2.17	1.40
11	27.66	31.93

6. おわりに

今回用いたモデルは，従来のパラメトリックなモデルと比較すると，モデルの自由度が大きく，カーネル関数としてデータに合うように自由な関数形を選べる特徴がある．一方，どのように最適なカーネル関数形や平滑定数の最適値を求めるかについて，決められた方法はなく，分析者の勘と経験に委ねられている．マーケティング変数が増えるに従い，最適化はますます困難になると予想されるため，今後の検討が必要であろう．

参考文献

[1] Abe, Makoto (1991): "A Marketing Mix Model Developed From Single Source Data : A Semi-parametric Approach," MIT Doctoral Thesis.

[2] Guadagni, Peter M, and John D.C. Little (1983): "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, vol.2, no.3, Summer, 203–238.