

想起集合を考慮した マーケットシェア予測モデル

中川 慶一郎, 土井 秀文

1. はじめに

本研究では、消費者が購買意思決定の際に選択の対象とする商品の集合である想起集合 (evoked set) に着目するが、その理由は以下の通りである。

- (1) インスタント・コーヒーの場合、スーパーマーケットでは一般に 20 種類程度のアイテムを扱っており、分量・形態が大きく異なるものも存在するため、全アイテムが全パネリストの想起集合に含まれるとはいえない。
- (2) パネリストの平均購入回数が年間 7~8 回であるのに対して 8 割のパネリストが 2~4 種類程度のアイテムしか購入しておらず、ブランドロイヤルあるいはバラエティ・シーキングといった特徴は顕著に現われていない。

本来、想起集合は各消費者に内在するものであり、外部からの観察は不可能である。そこで、本研究では、個別のアイテムについて想起集合に含まれる確率を推定すると同時に、想起集合に含まれたもとのアイテムを選択する確率を推定するモデルを提案する。次に、得られた選択確率より検証期間におけるパネリスト内のマーケットシェアの予測を行う。

2. 従来研究及び提案モデルの概要

従来のブランド選択モデルにおいて、想起集合が考慮されているものは数少ない。それは想起集合が各消

なかがわ けいいちろう NTTデータ通信(株) システム科学研究所 〒163-10 新宿区西新宿 3-7-1
どい ひでふみ NTTデータ通信(株) 国際部
〒135 江東区豊洲 3-3-3

費者に内在する漠然とした集合であるため、その測定が困難であることに起因する。

Fotheringham [1] は、この集合をファジィ集合と捉えて分析を行った。しかし、このアプローチはロジット・モデルといったブランド選択モデルの適用が困難であるという点で応用範囲が限定される。これに対し、守口 [2] は想起集合が過去の購買されたブランドより構成されるものとし、そのもとでロジット分析を行っている。このモデルでは、過去に購買実績のないブランドは、想起集合には入り得ない。また、全ての可能な想起集合のもとでの選択確率を算出するためにブランドの組合せ全てを考慮しなければならず、大規模な問題に対しては、サンプル数、計算量といった点で問題が生じる。

そこで本研究では、想起集合に含まれる確率と、想起集合に含まれたもとのアイテムが選択される確率を、各アイテム毎に個別に推定することにより対象アイテムの選択確率を算出する。これにより、効用の比較が行われた結果、アイテムが選択されるという理論的側面が損なわれるが、上述の問題が回避される。

3. 本研究の仮説

本研究では以下の仮説に基づいたモデルを提案する。

- (1) 各アイテムが想起集合に含まれる確率は、売価掛け率以外のマーケティング変数 (エンド陳列、チラシ) 及びパネリストと各アイテムの距離によって規定される。
- (2) 各アイテムの選択確率は、売価掛け率にのみ依存し、上記の変数は想起集合に含まれたもとの選択確率には影響を及ぼさない。

4. 提案モデル

4.1 各アイテムの選択確率

パネリスト j ($j = 1, 2, \dots, m$) が購入時点 t_j ($t_j = 1, 2, \dots, T_j$) においてアイテム i ($i = 1, 2, \dots, n$) を選択する確率 $P(C_{ijt_j})$ は、以下のように定式化される。

$$P(C_{ijt_j}) = P(E_{ijt_j})P(C_{ijt_j} | E_{ijt_j}) + P(E_{ijt_j}^C)P(C_{ijt_j} | E_{ijt_j}^C) \quad (1)$$

但し、 $P(E_{ijt_j})$ は購入時点 t_j においてアイテム i がパネリスト j の想起集合に含まれる確率である。

ここで、選択確率は想起集合に含まれたもとの選択される確率であるので、

$$P(E_{ijt_j}^C)P(C_{ijt_j} | E_{ijt_j}^C) = 0 \quad (2)$$

となる。

次に、 $P(E_{ijt_j})$ と $P(C_{ijt_j} | E_{ijt_j})$ であるが、それぞれ以下のようにロジスティック曲線に従うものとする。

$$P(E_{ijt_j}) = \frac{\exp(U_{ijt_j})}{1 + \exp(U_{ijt_j})} \quad (3)$$

$$P(C_{ijt_j} | E_{ijt_j}) = \frac{\exp(V_{ijt_j})}{1 + \exp(V_{ijt_j})} \quad (4)$$

ここで、 U_{ijt_j} と V_{ijt_j} はそれぞれ以下のものとする。

$$U_{ijt_j} = \sum_{i'=1}^n (\alpha_{i'i'}^{(1)} x_{i'jt_j} + \alpha_{i'i'}^{(2)} d_{i'j}) \quad (5)$$

$$V_{ijt_j} = \sum_{i'=1}^n \alpha_{i'i'}^{(3)} r_{i'jt_j} \quad (6)$$

$$\text{但し、} x_{i'jt_j} = \begin{cases} 0, & (\text{エンド無, チラシ無}) \\ 1, & (\text{エンド有, チラシ無}) \\ 2, & (\text{エンド有, チラシ有}) \end{cases}$$

d_{ij} : アイテム i とパネリスト j の距離

$r_{i'jt_j}$: 売価掛け率

$\alpha_{i'i'}^{(1)}, \alpha_{i'i'}^{(2)}, \alpha_{i'i'}^{(3)}$: パラメータ

また、パラメータは以下の制約を満たすものとする。

$$\sum_{i'=1}^n \{ (\alpha_{i'i'}^{(1)})^2 + (\alpha_{i'i'}^{(2)})^2 \} = 1 \quad (7)$$

$$\sum_{i'=1}^n (\alpha_{i'i'}^{(3)})^2 = 1 \quad (8)$$

4.2 各アイテムのシェア

各アイテムのパネリスト内でのシェアを予測するためには、各パネリストがインスタント・コーヒーを購入する頻度も同時に予測する必要がある。しかし、当コンペの検証データでは、パネリストの購入年月日、そのときのマーケティング変数は既知である。従って、アイテム i のシェアの期待値 MS_i は以下ようになる。

$$MS_i = \frac{\sum_{j=1}^m \sum_{t_j=1}^{T_j} P(C_{ijt_j})}{\sum_{j=1}^m T_j} \quad (9)$$

5. データ解析

5.1 アイテム、説明変数

表1は、各アイテムの販売量とマーケティング変数の平均の表である。ここで、アイテム5の販売量は他のアイテムに比べて少なく、1人のパネリストがその半数を購入しているため、「その他」として扱うこととする。

マーケティング変数については、年間を通じて変化のないものは除外する。また、アイテム i とパネリスト j の距離は、数量化理論 III 類により空間に付置した際のユークリッド距離を用いることとする(図1)。

表1: 販売量と変数の平均

アイテム	販売数量	売価掛率	エンド陳列	チラシ
1	798	0.852	0.175	0.011
2	661	0.905	0.176	0.008
3	619	0.820	0.227	0.055
4	581	0.823	0.207	0.040
5	74	1.000	0.000	0.000
6	468	0.946	0.129	0.026
7	299	0.800	0.295	0.026
8	203	0.879	0.006	0.000
9	173	0.915	0.000	0.000
10	147	1.000	0.000	0.000
その他	1601	0.690	0.000	0.000

(但し、エンド陳列及びチラシは有のとき1、無のとき0)

表 2: 実測値と予測値 (93 年)

アイテム	販売量		マーケットシェア	
	実測値	予測値	実測値	予測値
1	798	792.859	0.142	0.141
2	661	631.885	0.118	0.112
3	619	636.651	0.110	0.113
4	581	565.920	0.103	0.101
6	468	489.167	0.083	0.087
7	299	307.013	0.053	0.055
8	203	203.781	0.036	0.036
9	173	177.385	0.031	0.032
10	147	157.444	0.026	0.028
その他	1675	1661.896	0.298	0.296
合計	5624	5624.000	1.000	1.000

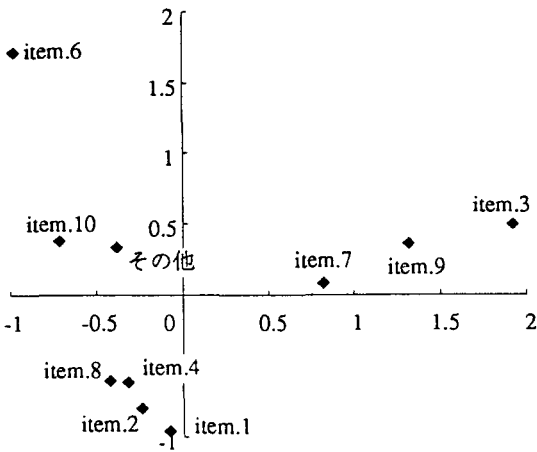


図 1: 数量化理論 III 類の結果

5.2 パラメータの推定

以下の式を最大化することによりパラメータの推定を行う。推定結果を表 5 に示す。

$$\prod_{j=1}^m \prod_{t_j=1}^{T_j} P(C_{ijt_j}) \quad (10)$$

5.3 予測結果

サンプルデータ、検証データに対する販売量及びマーケットシェアの予測値を表 2 と表 3 に示す。また、評価基準であるタイルの不一致係数 U を表 4 に記す。

6. おわりに

本研究では、想起集合に基づいたブランド選択行動をモデル化し、マーケット・シェアを予測するモデルを提案した。このモデルを利用することにより、アイテムが消費者の選択対象になっている度を定量化することが可能になるとともに、その際に競合アイテムが何であるかといったマーケティング上の示唆が得られることが期待される。

参考文献

- [1] Fortheringham, A. S.: "Consumer store choice and choice set definition," *Marketing Science*, 299-310, (1988).

- [2] 守口 剛: "想起集合を考慮したブランド選択モデル," 日本 OR 学会 1994 年度春季研究発表大会アブストラクト集, 75-76, (1994).

表 3: 実測値と予測値 (94 年)

アイテム	販売量		マーケットシェア	
	実測値	予測値	実測値	予測値
1	458	464.941	0.146	0.148
2	395	346.271	0.126	0.110
3	343	370.040	0.109	0.118
4	266	365.465	0.085	0.116
6	264	285.480	0.084	0.091
7	125	159.155	0.040	0.051
8	73	107.308	0.023	0.034
9	173	103.386	0.055	0.033
10	47	88.657	0.015	0.028
その他	1003	856.297	0.319	0.272
合計	3147	3147.000	1.000	1.000

表 4: タイルの不一致係数

	サンプルデータ	検証データ
不一致係数	0.0115	0.0842

表 5: パラメータの推定値

	アイテム	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 6$	$i = 7$	$i = 8$	$i = 9$	$i = 10$
$\alpha_{ii'}$ ⁽¹⁾	$i' = 1$	0.467	-0.246	-0.041	-0.089	-0.046	-0.078	-0.067	0.000	-0.018
	$i' = 2$	-0.197	0.426	-0.033	0.057	-0.030	-0.019	-0.088	-0.054	-0.043
	$i' = 3$	-0.040	-0.035	0.288	-0.050	-0.043	-0.152	-0.068	-0.097	-0.079
	$i' = 4$	-0.098	-0.245	-0.082	0.601	-0.022	-0.017	-0.094	-0.058	-0.092
	$i' = 6$	-0.028	-0.002	-0.014	0.036	0.041	-0.035	-0.076	-0.053	-0.046
	$i' = 7$	-0.026	-0.057	-0.098	0.025	-0.001	0.196	0.107	0.046	0.039
$\alpha_{ii'}$ ⁽²⁾	$i' = 1$	-0.784	-0.256	0.114	0.022	0.134	0.090	0.132	0.131	0.136
	$i' = 2$	-0.252	-0.737	0.145	-0.064	0.097	0.124	0.065	0.136	0.111
	$i' = 3$	0.031	0.116	-0.824	0.114	0.072	-0.117	0.138	-0.179	0.062
	$i' = 4$	-0.008	-0.082	0.157	-0.708	0.106	0.068	-0.128	0.104	0.042
	$i' = 6$	0.074	0.071	0.086	0.088	-0.953	0.067	0.114	0.062	-0.029
	$i' = 7$	-0.020	0.080	-0.153	-0.003	0.039	-0.921	0.042	-0.034	0.016
	$i' = 8$	-0.005	-0.075	0.041	-0.245	0.001	-0.017	-0.901	-0.027	-0.092
	$i' = 9$	-0.035	0.015	-0.281	0.009	-0.059	-0.104	-0.040	-0.922	-0.063
	$i' = 10$	-0.016	-0.044	-0.068	-0.127	-0.172	-0.086	-0.201	-0.108	-0.959
$i' = 11$	0.209	0.215	0.211	0.140	0.075	0.132	0.178	0.184	0.101	
$\alpha_{ii'}$ ⁽³⁾	$i' = 1$	-0.227	0.220	0.373	0.417	0.393	0.247	0.035	-0.033	0.216
	$i' = 2$	0.604	-0.728	0.387	0.178	0.404	0.343	0.141	0.290	0.318
	$i' = 3$	0.191	0.242	0.000	0.185	0.398	0.622	0.086	0.464	0.430
	$i' = 4$	0.463	-0.067	0.285	-0.664	0.381	0.262	0.599	0.454	0.309
	$i' = 6$	0.372	0.101	0.364	0.068	0.000	0.138	-0.052	0.148	0.313
	$i' = 7$	0.353	-0.186	0.393	0.168	0.279	-0.536	-0.194	-0.115	-0.052
	$i' = 8$	0.081	0.522	0.378	0.485	0.291	0.167	-0.751	-0.327	0.379
	$i' = 9$	0.040	0.191	0.367	0.194	0.384	0.101	0.054	-0.584	-0.507
	$i' = 11$	0.247	0.071	0.257	0.112	0.264	0.144	-0.075	0.107	0.266