

器具所有パターンによる

需要家分類とガス販売量予測

安井 誠 一

1. はじめに

現在、当社においてさまざまな形でガス器具の調査が行なわれている。その1つに総合器具調査とよばれるものが毎年実施されている。この調査は全需要家のなかから、約7万件をサンプリングし、個々の需要家の器具の所有台数や、過去1年間での購入台数、処分台数などを質問している。

本稿においてはこれらの情報をもとに、器具の所有パターンの分析、需要家の購買行動の分析、あるいは器具とガス使用量を結びつけた解析の結果を報告する。

われわれは、ガスはいくつかの器具の組み合わせで使用されているという考えに着目し、「所有器具」と「使用量」という2つを結びつけるために、器具の所有パターンというものを考えた。つまり需要家の器具の所有パターンには代表的なものがいくつかあり、個々の需要家は何年間かの間にそのパターンを推移し、それにともなってガス使用量も推移していくであろうという仮定を立ててモデルをつくっている。このモデルにおいてはかなり大胆な仮定をいくつかおいているが、計画段階の資料とするには十分に有効な数値がえられていると思われる。

2. クラスタ分析による需要家の分類

需要家の分類をするときに特性値として、普及率、代替性などを考慮してつぎのようなものを選んだ。

小型湯沸器、大型湯沸器、ガス炊飯器、
ガステーブル、一口コンロ、レンジ、ストーブ、
FFストーブ、風呂、オープン

以上10の特性値に関してクラスタ分析を行なう。各特性値に対しその器具をもっているか、いないかの(0, 1)型の変数で各個体(需要家)を記述し、個体間の距離をはかる尺度をつぎのようにきめる。いま2つの個体をベクトル表現で x, y とすると

$$x = (x_1, x_2, \dots, x_{10}), \quad y = (y_1, y_2, \dots, y_{10})$$

$$x_i, y_i = 0 \text{ or } 1$$

と書くことができる。このとき $x_i \neq y_i$ である特性値の個数を k とすると、2つの個体の距離を

$$d(x, y) = k/10$$

で定義する。この距離の考え方は「もち方の異なっている器具が多いほど2つの個体は離れている」という定義である。この距離に対していくつかの手法を適用してみた結果、最長距離法による分類がもっともうまく分類でき、各クラスターの特徴もわれわれが考えている代表的パターンに近いものがえられた。

今回は特性値として10種類の器具を選んでいますが、実際の器具の種類はもっと多く、これらのすべてについての解析をするためには数量化Ⅲ類などの方法によって、軸の縮約をしてからクラスタ分析をするほうが問題の見とおしはよくなるであろう。しかしこのような手法を適用すると、軸の解釈がうまくいかない場合などは、えられるクラスターに対する名前づけがむずかしくなり、かならずしもよい指標がえられるとはかぎらない。これらの点を考えると、どのような手法を適用してデータの解析をすればよいのかという判断を事前に行なうことは、かなりむずかしいものとなる。これらの解析の結果から器具の所有パターンの代表的なものとして以下の9つのパターンをあげてみた。

- ①小型湯沸器、炊飯器、ガステーブル、ストーブ、風呂
- ②小型湯沸器、一口コンロ、風呂
- ③一口コンロ
- ④大型湯沸器、一口コンロ、ガステーブル、ストーブ
- ⑤大型湯沸器、炊飯器、レンジ
- ⑥小型湯沸器、ガステーブル、ストーブ
- ⑦大型湯沸器、レンジ、風呂
- ⑧小型湯沸器、一口コンロ、炊飯器、ストーブ、風呂
- ⑨器具なし

使用したデータは50年度総合器具調査のなかから、家庭

表 1 パターン別分布と平均使用量

	①	②	③	④	⑤	⑥	⑦	⑧	⑨
既設需要家	0.383	0.184	0.227	0.031	0.022	0.032	0.010	0.037	0.074
新設需要家	0.213	0.153	0.196	0.025	0.022	0.029	0.015	0.022	0.325
平均使用量	1023	767	433	1406	1557	486	1990	1003	488

用需要家1570サンプルを抽出したものである。

第2段階の解析として、任意の需要家は前述した距離の意味で、これら9つのパターンのなかでいちばん近いものに属していると仮定して、各パターンの分布とその平均使用量を求めると表1のようになる。

表1からみてまず注目すべきことは①、②、③の3つのパターンが全体の80%を占めており、この3つが当社需要家の代表的なパターンといえよう。各パターンの特徴を総合的に判断すれば、多様化の傾向をもつパターンと単純化の傾向をもつパターンの2とおりがあると考えられる。これらのパターンの需要家のそれぞれにどれくらいの購買力があるのかということをするために、つぎにパターン間の推移を考えてみる。

3. パターン間の推移モデルと使用量予測

いま、全需要家に対して、どのパターンに属しているかという状態ベクトル $\pi(t)$ とその状態間の推移行列 P がわかっているとすると

$$\pi(t+1) = \pi(t) * P \quad (1)$$

の式から $t+1$ 時点での状態ベクトルがえられると仮定する。さらに各パターンでの平均使用量ベクトルを w とすると、需要家1件あたりの期待平均使用量 $G(t)$ は

$$G(t) = \pi(t) * w \quad (2)$$

で求まり、期待平均使用量増 $\Delta G(t)$ は

$$\Delta G(t) = G(t+1) - G(t) \quad (3)$$

で求まるものとする。ここで需要家件数を $C(t)$ とすると

期待全使用量増 $\Delta TG(t)$ は

$$\Delta TG(t) = C(t) * \Delta G(t) \quad (4)$$

で求まる。以上が推論の基本的考えである。ここで暗黙のうちに器具の所有パターンが推移すれば、ガスの使用量も推移していくであろうという仮定がもちこまれている。

つぎに推移行列 P を求める。情報として所有器具数、購入器具数、処分器具数がわかっており、これをもとに1年前の所有器具を(所有器具-購入器具+処分器具)として求めることができ、1年前のパターンを決定することができる。このようにして2時点に関するクロス表をつくり、パターン i からパターン j への推移の件数を f_{ij} とすると推移確率として $p_{ij} = f_{ij}/f_{i\cdot}$ を用いて P を決定する。

このモデルはマルコフモデルであり、ここでマルコフ性と定常性について若干の考察をしてみる。器具の所有パターンが直接的な要因として推移確率に影響するものとはあまり考えられない。ここでは間接的な意味でのマルコフ性の成立を仮定して問題を記述している。さらに定常性についてもそれを仮定しうほどの論拠はもっていない。われわれはこれらの仮定が現実のなかで成立しているか否かという点への検証について、現在のところそれほどこだわる必要がないのではないかと考えている。問題の見とおしをよくするために、非常に多くの状態をまとめていくつかの状態に集約してあつまっている。この意味からも厳密な意味でのマルコフ性が成りた

表 2 推移行列 P

50年 49年	①	②	③	④	⑤	⑥	⑦	⑧	⑨
①	0.9956			0.00185	0.00045		0.00005	0.00209	
②	0.07762	0.90631		0.00386		0.00579	0.00028	0.00615	
③	0.0076	0.00944	0.95751	0.00516	0.00066	0.00059	0.00022	0.01881	
④	0.06132	0.00061		0.92957	0.00061	0.00546	0.00061	0.00182	
⑤	0.07004		0.007	0.01868	0.89883		0.00311	0.00233	
⑥	0.07037		0.0033	0.0077		0.91864			
⑦	0.05714	0.04603	0.00952	0.01905	0.03175		0.83016	0.00635	
⑧	0.04995			0.0021	0.0021			0.94585	
⑨	0.15749	0.11264	0.0428	0.0111	0.01402	0.02016	0.00292	0.00541	0.63346

表 3 各時点での状態分布

	①	②	③	④	⑤	⑥	⑦	⑧	⑨
時点-1	0.417	0.178	0.22	0.033	0.022	0.032	0.008	0.042	0.047
2	0.446	0.169	0.213	0.035	0.021	0.032	0.007	0.046	0.029
3	0.472	0.159	0.206	0.036	0.02	0.031	0.006	0.05	0.019
4	0.496	0.149	0.198	0.037	0.019	0.03	0.006	0.053	0.012

たないところでマルコフモデルをあつかっていることになる。定常性についても経済的な要因などはかなり時間依存しているものと考えられよう。

したがって需要家の購買行動に影響すると思われる要因に対してすべて仮定の成立を検証する必要がある。解析をする以前にこれらの労力をはらうことにそれほど意味があるとは思えない。ある程度の仮定の崩れは無視して、解析した結果からモデルの適合性をはかる別の方法を見つけるほうがより実りのある結論がえられるであろう。厳密さを重視するとモデルが複雑になりすぎるきらいがある。われわれは単純なモデルで問題の本質をはずすことなく記述できればよいのであり、仮定の崩れの許容範囲のようなものがあれば十分であろう。社会における現象のモデル化においては、数学的厳密さよりはむしろ現実との適合性の評価基準が重要であろう。以上のような考え方に立ってマルコフ性の成立と短期的な定常性を仮定して解析をすすめていく。

いま求めた P と表 1 のパターンの分布を初期状態としてつぎのようにしてガス販売量の増加の予測式を求めた。

- $C_1(t)$: t 時点での既設需要家件数
 $C_2(t)$: " 新設 "
 $\pi(t)$: t 時点での既設需要家状態ベクトル
 π_0 : 新設需要家状態ベクトル
 $G_1(t)$: t 時点での既設需要家期待平均使用量
 G_2 : 新設需要家期待平均使用量
 w : パターン別平均使用量ベクトル (m^3 /件)
 $TG(t)$: t 時点での全ガス使用量 (m^3)

とすると t 時点から $(t+1)$ 時点の間でのガス使用量の増加は (2), (3) 式より

$$\begin{aligned}
 \Delta TG(t) &= C_1(t) * \Delta G_1(t) + C_2(t) * G_2 \\
 &= C_1(t) (\pi(t+1) - \pi(t)) * w \\
 &\quad + C_2(t) * \pi_0 * w \quad (5)
 \end{aligned}$$

により求まる。 $\pi(t)$, π_0 , w として表 1 の数値を用い、さらに P から $\pi(t+1)$ を求めると表 3 の数値がえられた。これらの数値を用いて (5) 式を書きなおすと

$$\Delta TG(t) = 19.23 \times C_1(t) + 714 \times C_2(t) \quad (6)$$

という式がえられる。右辺の第 1 項は既設需要家の所有

パターンの変化にともなう増分で、第 2 項は新設需要家の増加による増分である。 $C_1(t) = 500$ 万件、 $C_2(t) = 20$ 万件とすると $\Delta TG(t) \doteq 24000$ 万 m^3 となり、実績値にてらして考えてみるとかなりよい値と考えられる。

ここで (5) 式を考えてみよう。われわれはただ単にガス器具の販売台数を増やすことを目的とするのではなく、ガス販売量の増加に結びつく器具販売政策を模索している。(5) 式の $C_2(t)$ は新設需要家の獲得に対する営業政策に対して変化するものであり、これらは多大な導管投資を必要とし、供給計画とからみ合いただ単純に多くの需要家を獲得すればよいというものではない性質のものである。もう 1 つわれわれの営業努力によって制御可能な変数は $(\pi(t+1) - \pi(t))$ である。どのような器具をどのような層に売れば、もっともよくガス販売量の増加につながるか。あるいはどのような層の購買力が大きいのかという点に興味がある。この点についてなんらかの判断をするために推移行列 P について考察してみよう。

4. 器具販売政策に関する考察

パターンの推移の傾向をつかむために将来 5 年間の状態分布の表をながめてみる。全体的にながめてみると、パターン①の需要家の増加傾向がだんぜん多く、5 年後には約 50% を占めることになる。つぎにパターン⑧で、そのつぎがパターン④ということになっており、その他は減少している。①, ⑧は多様化傾向をもつ中所得層の需要家の代表パターンと考えられ、④は大型化傾向をもつ高所得層の代表パターンであると考えられる。要するに、全体の傾向として中所得層は①のパターンへ、高所得層は④のパターンへと推移していくようである。ここで 1 つ注意しなければならない点は、マルコフモデルにおいては、ある状態にいる需要家すべてが等しい推移確率をもっていると考えることである。実際にはその状態に滞在している時間、サンプリングによるひずみなどがあり、かならずしもすべての需要家が同一の過程で行動するとはかぎらないであろう。この意味からなんらかの修正をして結果をながめる必要がある。それではこれらのパターンへどのパターンから推移してくるのが多いのかということを考えてみよう。推移行列 P をみると①

のパターンへの推移の率がいちばん多いのはパターン⑨で15.7%の需要家が①に推移してくるわけであるが、この層の需要家は全体に対しての構成率が7.4%と低いのでそれほど大きな市場とは考えられない。しかしパターン⑨の需要家が他への推移がもっとも多いことは注目され、多くの器種に関して市場となりうることを示している。ここで注目すべきは②から①への推移であろう。このパターンは全体の構成率のなかで18.4%あり、①への推移が7.76%ある。この推移がおこるときに購入されると考える器具は炊飯器あるいはストーブと考えられる。したがって、パターン②の需要家はこの2器種に対する有望な市場と考えられ、営業努力によってガス販売量の増加が1件について256m³/年見込まれるであろう。①のパターンの需要家は99.56%と定着率が非常に高く、この9つの状態のなかの安定状態といえる。④のパターンに推移してくるのは、⑤と⑦の需要家の率が高いがこれは高所得層需要家の多様化傾向とみるべきであろう。さらに①から④への推移の率が1.85%と高い点も着目できる。この推移は需要家の所得水準の上昇によるものと思われ、必需品としてガス器具を購入したのではなく、生活の快適性、便利性などを重視して大型化傾向へと推移したのである。③の一口コンロだけという需要家についてみると、他パターンへの推移が少なくこれらの需要家の購買力はこれ以上望めないといえるであろう。

ここでもう1点いえることは、地域特性を考慮した販売政策のほうが合理的であろうという点である。いままでの考察は全体に対しての状態ベクトルを基本にして議論をすすめてきたが、実際問題としてはこの状態ベクトルは地域により異なっていると考えたほうが妥当であろう。たとえば一口コンロだけを所有している需要家の密集している地域で高級品などの販売努力をしてもそれほど売れるとは考えられず、むしろ小型湯沸器などの販売を重点的に行ない、③から②への推移確率をあげる努力をしたほうがより効果的といえるであろう。逆に④、⑤、⑦などの大型化の需要家の多い地域では、むしろFFストーブなどの高級品の販売に重点をおいて多様化を推進することも可能であろう。このように所有器具のパターンから需要家の潜在購買能力のようなものを、地域的に把握して器具販売計画を立てることが効果的であると考えられる。

5. 今後の方向性

現在、マーケティングへの適用を目的とした解析方法は、たとえば普及率、市場占有率といったような数値を求めて現状を把握しようとしている。これらは静的な側面をとらえるものであり、たしかに現状をつかむことは

可能であろう。しかし現実はいちばん動的であり、われわれはえられた結果を政策に反映するためには、その動的側面をとらえる必要がある。そのためには市場のなかに内在する変動性、あるいはその変動の構造をつかみ、それらの構造をみきわめたいうえで、的確な判断をする必要があろう。さらに問題をむずかしくしていることは、当社が器具の販売会社ではなく、ガス事業を営む公益企業であるという点であろう。われわれには安定供給の義務がありただ単に器具の販売台数を伸ばし、ガス販売量をふやせばよいというものではない。むしろ器具の販売台数やガス販売量の予測などからえられた情報と、安定供給の限界などとてらし合せたなかでの最適化をはかることに目的がある。これらのなかの1つの側面として器具販売とガス販売量という2つをとらえてモデルをつくったわけである。ガス販売量の予測ということについて考えてみれば、ほかに気温という要因が大きくきいてくる。さらに(5)の予測式についてみると、既設需要件数と新設需要件数の2つを説明変数とみて回帰モデルにあてはめることも可能であろう。しかしこれらのモデルの1つの欠点は、係数の意味がつかみえないことであり、具体的な政策に結びつきにくい。この点から、予測という観点に立てばかなり大きな要因を落したモデルといえるが、市場の構造を考えているという点で意味があるといえるであろう。

本稿における考察は、販売量の予測にしろ、器具の販売に関するものにしろ、全体の傾向を論ずるためのごくあらい議論であった。これらの結果をもとに実際の政策を立てるためには、さらにモデルを現実に近い努力をする必要がある。たとえばこのモデルのなかでは競合する製品——石油ストーブ、電気炊飯器など——への考察が欠けている。あるいは分類の手法に関する検討やマルコフモデルへの適用の妥当性などの検討も不十分であろう。

地域特性を考慮した解析なども行なえば意味のある結果がえられるであろう。これらの改良をくりかえし、出てきた結果やそのほかの情報をもとにモデルの現実への適合性を高めていくことが、われわれがこれからやるべき課題といえよう。

参 考 文 献

- 1) 小山昭雄：「マルコフ過程とその周辺」(東洋経済)
- 2) 関根智明訳：「ダイナミックプログラミングとマルコフ過程」(培風館)
- 3) 矢島敬二：「クラスター・アナリシス」(日科技連多変量解析研究会・MA研資—1972—022)
(やすい・せいいち 東京瓦斯システムセンター)