

顧客セグメントに基づくサービス利用形態の分析

1X14C105-1 原友美
指導教員 大野高裕

1. はじめに

近年、オンラインツールの拡大に伴い、企業の顧客への直接的なアプローチやコミュニケーションが盛んに行われるようになった。また、消費者の価値観の多様化により従来のマスマーケティングが通用しなくなっている。こうした背景から、顧客一人ひとりの趣味嗜好に合わせたダイレクトマーケティングが幅広い業界で取り込まれるようになった。特にスーパーマーケットなどの小売店においては、顧客の購買履歴データから、デモグラフィックな情報以外にも、購買行動についての情報を得ることが可能であり、実際に企業でも RFM 分析やデシル分析による顧客分析が導入されている。

一方、近年の少子高齢化や共働き世代の増加などを背景に、消費者のライフスタイルは多様化しており[1]、こうした顧客一人ひとりのライフスタイルの違いが日常の購買行動に表れていると考えられる。

そこで本研究では、顧客の購買行動について、購買の多様性と商品購買傾向に関する 2 つの概念に基づく指標を取り入れ、2 つの指標と購買行動の関係を明らかにする。また、上記の 2 つの指標により顧客をクラスタリングし、ライフスタイルに関する分類結果を導入したモデルの構築を行うことで、セグメント単位でのプロモーションの可能性について検証する。

2. 従来研究

2.1. 新美ら[2]

新美ら[2]では、Deep Neural Network を用いて、オムニチャネルにおける顧客行動の「多様性」を考慮した場合に、将来の購買行動の予測精度が向上する可能性について検証している。顧客の行動の多様性とは、顧客が Web 上で閲覧または購買した商品ジャンルの幅広さや特定のジャンルへの依存度のことである。顧客 k の購買された商品カテゴリの多様性は、経済指標である HHI (ハーフィンダール・ハーシュマン指数) を用いて、

$$HHI_k = \sum_{i=1}^{n_k} s_{ik}^2 \quad (1)$$

と表している。ただし、 n_k は顧客 k の購買した商品カテゴリ数、 s_{ik} は当該顧客 k の購買全体に占めるカテゴリ i の割合である。顧客の購買が特定のカテゴリに依存している場合に HHI_k の値が大きくなり、カテゴリごとのシェアの偏りが無い場合に HHI_k の値は小さくなる。

2.2. 従来研究の問題点

新美ら[2]では、多様性変数を導入した場合は、導入していない場合に比べて予測精度が向上していることから、多様性が購買行動を説明する要因となっていると言える。しかし、具体的に多様性が購買行動に対してどのような影響を与えているかについては明らかにされていないため、企業のダイレクトマーケティングに活かすことが難しいという問題点がある。

3. 提案モデル

3.1. 概要

本研究では、顧客のライフスタイルを説明する変数として、購買金額の多様性と、最大金額カテゴリの 2 つの指標

を用意した。これらを説明変数として、購買金額を目的変数とする購買行動モデルが妥当であるかを検討する。

次に、購買金額の多様性と最大金額カテゴリの 2 つの指標が総購買金額に影響を与えることを確認できたならば、この 2 つの指標を用いて自己組織化マップによる顧客のクラスタリングを行い、その結果からライフスタイルについて考察を行う。最後に、クラスタリング結果を顧客のライフスタイルに関するセグメント属性として導入し、新たに購買行動モデルを構築し、妥当性を検討する。

3.2. 購買金額の多様性について

本研究では、顧客の多様性を表す指標として、購買金額の多様性を用意した。これは、商品をカテゴリに分類した際の、商品カテゴリ間の購買金額のばらつきを表したものであり、新美ら[2]の研究で用いられた HHI を参考に、次のように定義した。

$$Diversity_k = \frac{1}{\sum_{j=1}^n \left(\frac{s_{kj}}{s_k} \right)^2} \quad (2)$$

ただし、 n は店舗の商品カテゴリ数、 s_{kj} は顧客 k のカテゴリ j における総購買金額、 s_k は顧客 k の商品全体の購買金額を表している。この購買金額の多様性は、新美ら[2]の HHI の逆数を取っているため、 $Diversity_k$ の値が大きければ購買金額の多様性が大きく、あらゆる商品カテゴリで購買が行われていると解釈することができる。

3.3. 最大金額カテゴリについて

購買金額の多様性を表す $Diversity_k$ では、商品カテゴリ間の購買金額の偏り具合を定量的に表すことができる。しかし、具体的にどのカテゴリを重点的に購買しているのかを表現することができない。そこで、本研究では顧客の商品購買傾向を表す指標として、最大金額カテゴリを定義する。これは「どの商品カテゴリに最もお金をかけているか」を表す指標で、顧客の購買金額のカテゴリ構成比率が最大となる商品カテゴリを求めることで与えられる質の変数である。

3.4. 顧客のクラスタリング

本研究では、顧客の多様性を考慮したセグメントに基づく購買行動の分析を行うために、顧客のセグメント分類を行う。顧客を購買金額の多様性と最大金額カテゴリの 2 つの指標によって顧客をクラスタリングする際の手法として、自己組織化マップを用いる。自己組織化マップとは、Kohonen が提唱した機械学習手法で、教師データを持たない学習型のニューラルネットワークモデルである。大規模データに対して効率的に結果を得られることやグルーピングにより新たな発見が可能であることなどを強みとしている。本研究ではどのような顧客セグメントが存在しているか未知の状態であることから、自己組織化マップの利用が有効であると考えられる。

3.5. 購買行動モデル

3.5.1. モデル 1

顧客を購買金額の多様性と最大金額カテゴリによってセグメントを分類するにあたり、これら 2 つの指標が総購買金額に有意性をもつ変数であることを確認する必要がある。そこで、購買金額の多様性と最大金額カテゴリを導

入した顧客の購買行動を、重回帰分析を用いて、

$$\text{Monetary} = b_0 + b_1 \text{Diversity} + \sum_{i=1}^n c_i x_i \quad (3)$$

と定式化する。なお、*Monetary*は総購買金額、*Diversity*は購買金額の多様性、 x_i は商品カテゴリ*i*のうち最大金額カテゴリを表すダミー変数である。また、 b_0 は定数項、 b_1 、 c_i は偏回帰係数である。

3.5.2. モデル 2

モデル 1 により購買金額の多様性と最大金額カテゴリの 2 つの指標が総購買金額を説明する変数として有意であることが示された後に、2 つの指標によって顧客をクラスタリングする。このライフスタイルに関する顧客のクラスタリング結果を導入して、購買行動モデルを

$$\text{Monetary} = b_0 + \sum_{l=1}^m d_l \text{group}_l \quad (4)$$

と定式化する。 group_l は分類された顧客の所属セグメント *l* を表すダミー変数であり、 d_l は偏回帰係数を表している。

4. 検証

本研究では、スーパーマーケットにおける ID 付き POS データを用いる。サンプルは 2001 年 1~6 月に利用した顧客 2300 人である。また、本研究で使用する商品のカテゴリは、商品 ID から売り場による分類を行い、チルド売り場商品をカテゴリ 1、常温売り場商品をカテゴリ 2、冷凍売り場商品をカテゴリ 3、消耗品をカテゴリ 4 とした。顧客全体の総購買金額の平均は 65,092 円であった。

購買金額の多様性と最大金額カテゴリを組み込んだ購買行動モデル 1 の構築結果を表 1 に示す。また、クラスタリングにより得られた各セグメントの特徴と総購買金額の平均値、分散を表 2 に、それに基づいた顧客の所属セグメントを用いた購買行動モデル 2 の構築結果を表 3 に示す。

表 1. モデル 1 構築結果

	標準偏回帰係数	Pr(> t)	
切片	0.000	1.000	
Diversity	0.283	0.000	***
x_1	0.119	0.038	*
x_2	0.159	0.006	**
x_3	-0.055	0.012	*

表 2. クラスタリング結果

	Diversity	x_1	x_2	x_3	人数	総購買金額の平均
G1	2.31	1.00	0.00	0.00	723	64596
G2	2.45	0.00	1.00	0.00	10	69913
G3	1.87	0.00	1.00	0.00	1511	65584
G4	2.99	1.00	0.00	0.00	1	5590
G5	2.90	0.00	0.00	0.15	53	57222
G6	3.00	0.00	1.00	0.00	2	20878

表 3. モデル 2 構築結果

	標準偏回帰係数	Pr(> t)
切片	0.000	1.000
group ₁	0.355	0.280
group ₂	0.072	0.160
group ₃	0.371	0.270
group ₄	-0.006	0.830
group ₅	0.095	0.380

5. 考察

モデル 1 では、導入した変数がどれも有意水準 5% を満たしたことから、顧客の購買金額の多様性および最大金額カテゴリは総購買金額を説明する要因として有意であると言える。また、標準偏回帰係数に着目すると、*Diversity*の係数が正の値を取ることから、顧客の購買金額の多様性は総購買金額に正の影響を与えると考えられる。*Diversity*の値が大きいとは、顧客が 1 つの店舗で複数の売り場の商品を満遍なく購買しているということであり、顧客が生活上の購買の多くを当該店舗に頼っていると捉えることができる。このことから、店舗内で幅広い購買を行う顧客ほど、総購買金額も大きくなると考えられる。

また、最大金額カテゴリのダミー変数の標準偏回帰係数に注目すると、カテゴリ 4 を基準として、カテゴリ 1、2 は正、カテゴリ 3 は負の値となった。カテゴリ 3 については、冷凍売り場商品を重点的に購買している顧客は普段から調理をする習慣が乏しく、当該店舗以外にもコンビニや外食で食事を用意しているために店舗での購買金額が小さくなる傾向となったと考えられる。一方で、カテゴリ 1 のチルド食品やカテゴリ 2 の常温売り場については、普段から調理の食材購入を目的として当該店舗を利用している顧客ほど利用頻度が増加したり一回あたりの購買金額が高くなるため総購買金額に正の影響が出ていると考えられる。

顧客のクラスタリング結果は、*Diversity*の値の大きさや最大金額カテゴリによって 6 つのグループに全顧客が分類されたが、グループによって人数に大きな差が出た。 x_2 の値が 1.00 を示したもののうち、所属人数の少ない G6 を除く G2、G3 の総購買金額の平均値は顧客全体の平均値を上回った。これはモデル 1 の構築結果において、最大金額カテゴリが 2 の場合に総購買金額に正の影響が出ることと一致した。

しかし、クラスタリング結果を用いたモデル 2 ではどの変数も有意水準を満たさなかった。顧客の所属セグメントが総購買金額を説明できなかった原因としては、クラスタリングの段階でグループ間の人数に大きな偏りが出たことから、最大金額カテゴリに関わらず、多様性が大きければ総購買金額は大きくなることが窺える。これにより、企業は顧客の多様性を向上させるために、店舗内のレイアウトの改善や、同時購買を誘引させるようなプロモーションが有効であると考えられる。

6. 結論と今後の課題

本研究では、購買履歴データから顧客の購買金額の多様性と最大金額カテゴリが総購買金額に与える影響を明らかにした。その結果、顧客の購買行動を、ライフスタイルの視点から考察することができた。一方で、上記の 2 つの指標による顧客のセグメント分類結果は総購買金額に対する有意性を示さなかった。

今後の課題としては、本研究で用いた指標に加え、顧客のデモグラフィック属性を考慮したセグメンテーションにより、総購買金額に有効なセグメントを決定することが挙げられる。

参考文献

- [1] 木島豊希: “2020 年のスーパーマーケット業界の課題と展望に関する調査研究”, 流通情報, Vol.43, No. 5, pp.40-58 (2012)
- [2] 新美潤一郎, 星野崇宏: “顧客行動の多様性変数を利用した購買行動の予測”, 人工知能学会, No. 32, pp. 1-9 (2017)