

# スケールの階層性から探る スーパーマーケットの消費者行動

金子 雄太, 矢田 勝俊

スーパーマーケットの顧客動線データと購買履歴データを使い、非集計データの分析モデルとして顧客の回遊行動に焦点を当てた購買行動モデル、集計データの分析モデルとして状態空間モデルのそれぞれの研究事例を紹介する。ここで強調するのはスケールの概念であり、非集計・集計で別々に考えられてきた分析手法に共通の視点を見いだすことである。回遊行動の自己相似性の概念からフラクタル次元を導入し、本指標が顧客の購買に与える影響を調べる。また、顧客の売場間の移動を含む潜在的な売場構造を状態空間モデルから推定し、店舗経営の改善策を模索する。

キーワード：スケール、フラクタル、2項ロジットモデル、状態空間モデル、顧客動線データ

## 1. はじめに

本稿ではスケール階層というキーワードを軸に消費者行動について考えていきたい。ある現象をわれわれが数理的に記述するとき、現象を相対化するための物差しが必要になる。身長を測定しようとするときは、〈cm (センチメートル)〉という単位で大きさを測るのが一般的であろう。われわれは〈cm〉という一つの尺度で個人間の身長の違いを相対化し、明確化しているのである。マーケティングの分野で重要な分析対象の一つは企業の収益である。収益を相対化する尺度は、日本では〈円〉、アメリカでは〈ドル〉である。このような物事の大きさを測るための尺度あるいは単位を一般的に“スケール”と呼ぶ。自然科学においては、スケールの混在がしばしば主要な研究テーマとなりうる。「蝶が羽ばたくと遠方で竜巻が発生する。」こんな話を耳にしたことはないだろうか。カオスという初期値に依存する予測困難な現象を「バタフライ効果」と表現した喩え話である。蝶の羽ばたきと竜巻との膨大なエネルギー差を考慮すれば荒唐無稽な話である。しかし両者の構図を俯瞰してみると、巨大な竜巻と小さな蝶の間にはスケール階層の構造が浮かび上がってくる。われわれの身体を例に考えてみよう。人体は分子、細胞、

組織などから構成される。分子は最もミクロなスケールであり、筋肉や骨、血液といった組織はマクロなスケールに属する。身体はこれらの部分的な要素から構成されると解釈するならば、身体を系 (system) と定義し、系の中には小さなスケールの部分系 (sub-system) が存在するとみなすことができる。これがスケールの階層構造である。

われわれを取り巻く世界は、さまざまなスケールの系が混在し、時にはそれらの間を越境する現象が観測される。銀河の渦巻き模様や太陽の爆発的なフレアの現象、生物の高度な生体機能、等々である。スケールの尺度を設定し部分系の連関を想定することで、線形システムでは観測できない“非線形”な現象の一端から消費者行動を探るのが、本稿で試みるアプローチである。

## 2. 分析データについて

本稿でわれわれは消費者行動にスケール階層の構造を見だし、マーケティングへの応用の可能性を模索する。考察するのは小売店内の消費者行動である。マーケティングでは、構築した理論は観測データによる証拠を取り込まなければ、実践的な手段とはなりえない。本稿で扱うデータは主に顧客動線データと ID 付 POS データである。ID 付 POS データを利用することで購買履歴と顧客属性を紐づけることができる。さらに顧客動線データを組み合わせれば、顧客の商品購入プロセスを明らかにすることができる。購買履歴、顧客属性、購入プロセスを組み合わせることで顧客理解の深化を目指すのが顧客動線研究である [1]。

### 2.1 パスデータ

Hui らは、顧客動線データを含むパスデータ (path

かねこ ゆうた  
関西大学データサイエンス研究センター  
〒 564-8680 大阪府吹田市山手町 3-3-35  
r148032@kansai-u.ac.jp  
やだ かつとし  
関西大学商学部  
〒 564-8680 大阪府吹田市山手町 3-3-35  
yada@kansai-u.ac.jp

表 1 顧客動線データの例

カート番号	時刻	$x$	$y$	エリア
101	10:10:31	55	339	V
101	10:10:32	55	341	V
⋮	⋮	⋮	⋮	⋮
101	10:15:42	148	195	F
101	10:15:43	149	179	F
101	10:15:44	150	178	F
⋮	⋮	⋮	⋮	⋮

data) は、物理的あるいはシミュレートされた環境における意識的なエージェントの行動履歴のことである、と述べている [2]. 言い換えると、次の three-tuple の  $P$  によってパスは与えられる.

$$P = \{S, A, X_A(t)\} \quad (1)$$

ここで  $S$  は環境,  $A$  はエージェント,  $X_A$  はエージェントの行動である.  $X_A$  は時刻  $t$  に依存する. こうした表現はエージェントシステムのシミュレーションなどに造詣がある読者にはわかりやすいだろう.  $X_A$  が時空間データのような形で観測されているとしても、環境やエージェントについて情報がなければ、それはパスデータとは呼ばない. パスデータはたとえば環境が物理的か非物理的か、連続か不連続か、といった基準で細分化していくことができる. 興味のある読者は Hui らの研究 [2] をご覧いただきたい. スーパーマーケットの顧客動線データの場合、環境は物理的であるが不連続として扱うのが通例である. マーケティングでは取り扱いの簡便さから、店内レイアウトを各売場単位の不連続な領域に分割し、その領域ごとに集計する方法が採用されてきた. このような領域区切りのレイアウトは、離散データを扱えるアルゴリズムを適用可能にし、複雑なデータをシンプルに表現できるため有効である. 一方で、売場単位で集計されたデータを用いると、モデルの精度は店内レイアウトの分割方法に大きく影響を受ける. この点を克服するため、レイアウトを連続的に扱うように、ノンパラメトリックな推定法であるカーネル密度推定法の導入が提案されている [3-5].

## 2.2 顧客動線データ

本稿で扱う顧客動線データは 2012 年の 9 月と 10 月に関東地方のスーパーマーケットで取得されたもので、関西大学データサイエンス研究センターから提供を受けたものである. 顧客動線データとは、IC タグの付いたショッピングカートの位置情報を時刻とともに記録したものであり、売場を 2 次元平面に展開すると座標

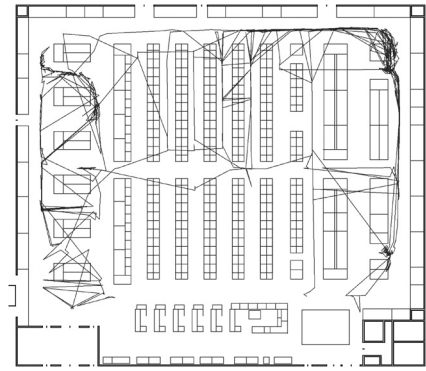


図 1 店内顧客動線の可視化

上の一点に位置情報が記録される. 表 1 では左から順にカート番号, 記録された時刻, カートの  $xy$  座標, 売場のエリアが示されている. カートの位置情報は秒単位で記録されており, 本稿で紹介する分析事例では,  $685 \times 654$  ピクセルのイメージデータに変換されている. 表中の V という記号は青果, F は鮮魚の売場を表している. 店内の外側エリアには青果, 鮮魚, 総菜, 精肉, 日配などの売場が配置されており, 内側エリアには一般, 菓子や酒類などの売場が並ぶ. 図 1 は店内レイアウトに一人の顧客動線を描いた可視化のマップである. タイムスタンプごとのカートの位置を直線で結んでいる. 図 1 の顧客の事例は典型的なもので, 顧客は店に入ると多くが外側エリアを訪問し, 買い物するとそのままキャッシュアウトしていく. 中には内側エリアも頻繁に訪問する顧客がいるが, 一般に内側エリアの訪問は外側エリアの訪問に比べると少ない. 顧客動線を店内レイアウトと一緒に可視化することで, 数理モデルに馴染みのない店舗管理者でも容易に顧客の購買プロセスを理解することができる. このような可視化の手法は, 実務の観点からは特に重要だと言える.

## 3. 非集計データの分析—マイクロな階層—

### 3.1 顧客動線のフラクタル次元

顧客動線を特徴づけるスケールは何であろうか. また, スケールの尺度を変えることで, 顧客の行動に違いが見られるだろうか. このことをフラクタルの概念を軸により深く考えていく. フラクタルとは一般的に, 自然界に存在する相似な構造をもつ物体および現象のことを意味する [6-8]. フラクタルは特徴的なスケールをもたず複雑な構造が内包されることで知られる. 鳥や湾の海岸に複雑に入り組んだ構造を見ることができよう. 海岸を遠くから眺めた場合と近くから眺めた場合で, 両者が似た形に見える錯覚を起こすことが

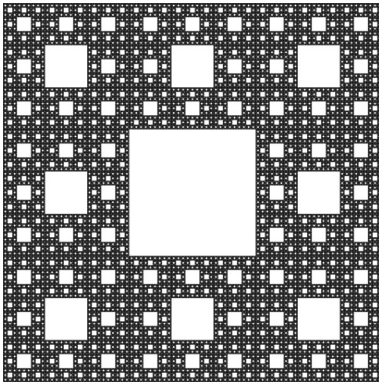


図2 シェルピンスキー・カーペット

ある。また、岩石を砕いた小片が元の石に似た形になるなど、複数の相似構造が見て取れるのがフラクタルの特徴である。図2は数理的な法則によって生成した、シェルピンスキーのカーペットと呼ばれるフラクタル図形である。四角形の内側に無数の相似な構造の四角形が内包されているのが確認できるだろう。

フラクタル図形はスケール不変性という重要な性質を備えている。ここで図2のフラクタル図形のような、2次元平面上に描かれた曲線を考える。この曲線を、一辺の基準長さを $\varepsilon$ とした矩形で被覆したとき、必要なセルの数を $N$ とする。比例係数 $c$ を用いると、 $N$ と $\varepsilon$ の間には以下の関係が成り立つ。

$$N = c\varepsilon^{-p} \quad (2)$$

ここで $p$ は正の実数である。この指数 $p$ をフラクタル次元と呼ぶ。式(2)のべき乗則で、 $\varepsilon$ をスケール変換すると

$$N(k\varepsilon) = c(k\varepsilon)^{-p} = k^{-p}N(\varepsilon) \quad (3)$$

となる。すなわち、 $N(k\varepsilon)$ と $N(\varepsilon)$ の間には線形の関係が成り立つために、自己相似の構造が生じていることがわかる。一般化すると、ある図形の次元とは、スケール長 $\varepsilon$ のセルによって図形を覆い尽くすために必要なセルの数 $N$ と $\varepsilon$ の関係から

$$D = -\lim_{\varepsilon \rightarrow 0} \frac{\ln N}{\ln \varepsilon} \quad (4)$$

により定められる。フラクタル次元は、フラクタル(半端な)という言葉の意味が示唆するように、通常の図形の次元が非整数次元に拡張された概念とも言い換えることができる。シェルピンスキー・カーペットのフラクタル次元は $p = \ln 8 / \ln 3 \approx 1.893$ であり、直線の次元1と平面の次元2からのずれが複雑な構造の特徴を

表2 フラクタル現象とフラクタル次元

例	フラクタル次元
銀河の星の空間分布	1.2
コッホ曲線	1.26
墨流し	1.3
雲の形	1.35
シェルピンスキー・ガスカート	1.58
シェルピンスキー・カーペット	1.89
ランダム・ウォーク	2

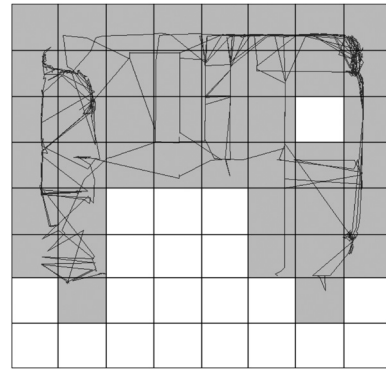


図3 顧客動線をセルで覆っていく様子

表現している。フラクタル次元とはスケールの尺度を変化させたときの、物体や現象の自己相似性を定量化する指標にほかならない。表2に種々の物体や現象のフラクタル次元を載せる。ここでは比較のための顧客動線は平面に描けるものとして定義しているため<sup>1</sup>、文献[7, 8, 10]を参考に次元数2までの例を載せている。

フラクタル次元を導出する方法はいくつか知られているが、ここではボックスカウント法と呼ばれる方法を採用する。ボックスカウント法は離散化された観測データへの適用に有効である。すなわち、スケール長 $\varepsilon$ のピクセルで曲線を覆っていき、動線部分を被覆するセルの数を数えることでフラクタル次元を求める[10]。図3は図1の顧客動線をセルで覆っていく様子を图示したものである。図3の場合、動線を含むダークカラーのセル数は43となる。顧客動線は $685 \times 654$ ピクセルの画像データとして表現されているため、スケール長を342ピクセルから最小の1ピクセルの単位まで変化させていき、顧客動線を被覆するセルの数を数えていく。被覆に必要なセルの数 $N$ とセルのサイズ $\varepsilon$ は図4で図示される関係にある。両対数表示で右下りの関係性が示唆される。この直線の傾きがフラクタル次元である。最小二乗法によって直線の傾きを決定すれば、フラク

<sup>1</sup> 時間軸を加えて3次元時空で顧客動線を定義することもできる[9]。

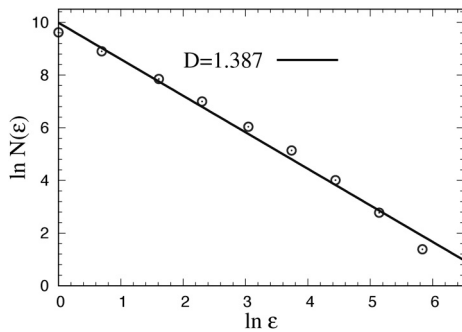


図4 セルの一辺の基準長  $\varepsilon$  と被覆数  $N$  との関係。直線の傾きは最小二乗法によって推定できる。調整済み決定係数は  $R^2 = 0.987$ 。

タル次元を 1.387 と推定することができる。このように比較的簡易な方法によって、顧客動線の自己相似性を評価することができる。顧客動線の自己相似性とは、マーケティングでは顧客の店内における回遊性あるいは売場訪問のカバー率を示していると解釈することができるだろう。特定のセルサイズでこれら进行评估すると、その結果はセルのサイズ（スケール）に依存してしまう。そこでスケール長を変えても不変な指標（ベキ乗則）を見いだすことで、観測に対してより強固な基準を与えることができる。フラクタルは自己再生可能である。ノイズ成分があっても、それ以外の直線部から直線勾配と切片を計算できるため、真のフラクタル次元を自己相似性の性質から再生できる。ノイズ環境下でも指標を定量的に評価できることは利用上重要である [8]。

### 3.2 2項ロジットモデルによる購買行動の分析

これまで顧客動線を測るスケールの尺度を変化させることで、動線の自己相似性という重要な性質を指標化することができた。次にマーケティングにおいて、フラクタル次元がいかなる意味をもつのかを明らかにしなければならない。顧客の購買行動への影響を考察するために2項ロジットモデルを導入する [9, 11]。  $i$  番目の人があるカテゴリーの商品を購入するかどうかを考えると、  $i$  番目の人が購買する確率  $Pr(y_i)$  は次のモデルで表現することができる。

$$Pr(y_i = 1) = \frac{\exp(\alpha_1 + \alpha_2 \cdot D_i)}{1 + \exp(\alpha_1 + \alpha_2 \cdot D_i)}, \quad (5)$$

ここで変数は以下のように定義される。

$$y_i = \begin{cases} 1: i \text{ 番目の顧客が購入} \\ 0: \text{それ以外} \end{cases}$$

$D_i$ :  $i$  番目の顧客のフラクタル次元

$\alpha_1, \alpha_2$ : パラメータ・ベクトルの成分

表3 パラメータの推定結果

Category	$\alpha_1$ and $\alpha_2$	Std.Error	$z$ -value
青果	-5.056	1.468	$z = -3.445^{**}$
	5.051	1.205	$z = 4.192^{**}$
鮮魚	-5.149	1.321	$z = -3.898^{**}$
	4.376	1.078	$z = 4.059^{**}$
精肉	-4.575	1.319	$z = -3.468^{**}$
	3.638	1.075	$z = 3.386^{**}$
惣菜	-1.849	1.297	$z = -1.425$
	1.352	1.057	$z = 1.278$
一般 B	-1.155	1.428	$z = -0.809$
	1.745	1.167	$z = 1.495$

(\*\* $p < .01$ )

モデルのパラメータ・ベクトルの推定結果を表3に示す。表中の上段が  $\alpha_1$ 、下段が  $\alpha_2$  の推定結果である。五つの商品カテゴリーの中で、青果、鮮魚、精肉の購買で統計的にフラクタル次元の重要性が高くなっている ( $p < .01$ )。逆に惣菜、一般 B のカテゴリーではフラクタル次元は購買に対して重要ではなかった ( $p > .10$ )。何を買うか迷って売場を回遊すると顧客動線は入り組み、フラクタル次元は高くなると考えられる。青果、鮮魚、精肉についてはフラクタル次元が購買に対し正の効果をもつために、回遊の複雑さやランダム性が強まると、顧客は商品を購入しやすくなると言える。野菜や魚、肉は材料となる素材であり、惣菜や一般 B の商品は加工品という特徴がある。素材は日々の値段の変動が激しく、また鮮度の影響を受けやすい。商品比較で考慮すべき要因が増えることが、顧客の店内回遊を促進し購買結果に結びついていると考えられる。また顧客層を考えると、素材の品をよく購入するのはファミリー層の主婦である傾向が強い。惣菜などの加工品は単身者が手軽さを求めて購入するだろう。青果、鮮魚、精肉の売場では主婦の店内回遊を促進するように、商品の陳列や告知を工夫するといった経営施策が有効だと考えられる。

このように具体的な事例分析で、フラクタル次元が経営施策のための指標として有効であることが明らかとなった。われわれはフラクタル次元を通して顧客動線の自己相似性の概念を浮彫りにした。この指標は小売マーケティングにおいて顧客の売場滞在時間や訪問数を統合した指標として利用でき、店頭プロモーションや商品配置、売場レイアウトの変更の効果を測定するのに有効である [12]。

## 4. 集計データの分析—マクロな階層—

3節では顧客一人の単位で回遊性と購買について考えた。本節では顧客ごとのマイクロな視点からいったん

離れ、売場全体を対象にするようなマクロな視点で消費者行動と購買について考えていきたい。マーケティングでは分析に集計データを利用することが主流である。個々には定まった特徴が見えなくとも、集計すればある一定の傾向が見えてくることが多いからである。集計データには顧客ごとの集計、売場ごとの集計、時間ごとの集計などさまざまな単位つまりスケールを導入することができる。ここでは1時間ごとの時間スケールで売場間の顧客の流れや移動の概念を状態空間モデルに組み込み、スーパーマーケットの売上分析に応用した事例研究 [13] を紹介する。

#### 4.1 状態空間モデルによる売場構造の推定

売上および各売場の訪問数を観測データとして、顧客の流動を考慮した売場の潜在的な構造をベイズ統計学で推定する。分析モデルは形式的には状態空間モデルと呼ばれる時系列モデルの形態をとる。状態空間モデルはマーケティングにおいて、市場の動的な変化を捉えるために、さまざまな領域で用いられてきた [14-18]。顧客動線データは空間情報を含むために、時空間的な扱いを可能にするフレームワークを構築しなければならない。状態空間モデルの観測変数を売場ごとに区切られた各商品カテゴリーの売上  $Y_t$  と顧客の売場訪問数  $N_t$  と定める。添え字  $t$  はこれらの変数が動的に変化していくことを意味する。変数  $Y_t$  と  $N_t$  は1時間ごとに、またカテゴリー別に集計された観測データである。 $Y_t$  と  $N_t$  をそれぞれ潜在変数  $\mu_{t,j}$ ,  $s_{t,j}$ ,  $r_{t,j}$  を用いて以下のように書き下す：

$$N_{t,j} \sim \text{Poisson}(e^{r_{t,j}}) \quad (6)$$

$$Y_{t,j} = \mu_{t,j} + s_{t,j} + r_{t,j} + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2) \quad (7)$$

添え字の  $j$  はカテゴリー別に考慮することを意味する。 $N_{t,j}$  は  $r_{t,j}$  をパラメータとするポアソン分布に従う。確率分布のパラメータが指数関数の形式で書かれるのは、ゼロ以上の値に変換するためである。売上  $Y_{t,j}$  は変数  $\mu_{t,j}$ ,  $s_{t,j}$ ,  $r_{t,j}$  の線形結合で表すことができる。 $\mu_{t,j}$ ,  $s_{t,j}$  はそれぞれトレンド項、季節調整項である。トレンド成分は2階差分のトレンドモデルに従うと仮定する [19-21]。

$$\mu_{t,j} = 2\mu_{t-1,j} - \mu_{t-2,j} + \delta_t, \quad \delta_t \sim N(0, \sigma_\delta^2) \quad (8)$$

トレンド項は長期的な時系列の変動を捉えるものとして導入され、ここでの  $\mu_{t,j}$  は売上  $Y_{t,j}$  の直接観測されない長期的な変動を表している。本分析の顧客動線データは、9月と10月の月単位のデータであり、1日を10時から18時まで1時間ごとに八つの時間帯に

表4 売場の隣接関係

Area No.	Category	隣接関係
1	青果, 鮮魚	2, 4, 5
2	惣菜	1, 3, 4
3	日配, 冷食, ドリンク, 精肉	2, 4, 5
4	一般 A, 一般 B	1, 2, 3, 5
5	一般 A, 菓子, 酒類, 中央通路	1, 3, 4

分割する。売上と訪問数は1時間ごとに記録されている。1日の変動は次の季節調整モデルに従うと仮定する [19-21]。

$$s_{t,j} = - \sum_{i=1}^7 s_{t-i,j} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \quad (9)$$

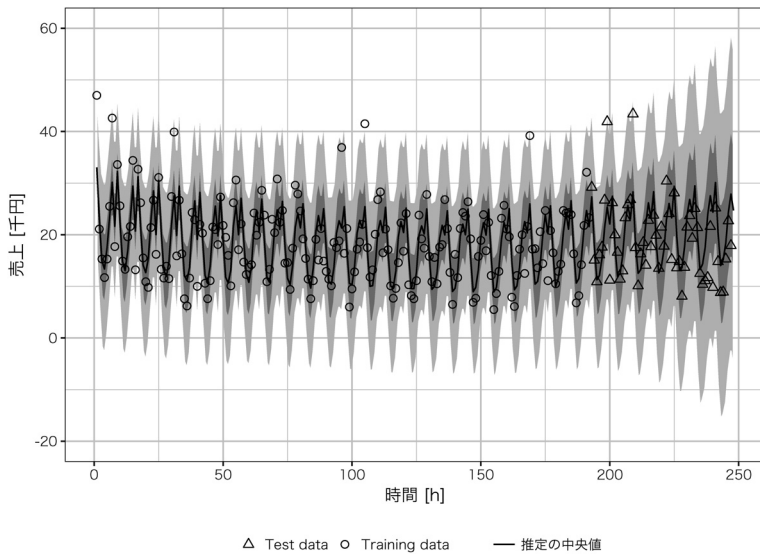
このように時系列の変動をトレンド項と季節調整項に分解することができる。さらに金子らの研究 [13] では、顧客の売場訪問数の効果がマルコフ確率場で表せるとして、これを状態空間モデルに導入した。訪問数の効果は売場に固有の性質をもつと解釈され、次のマルコフ確率場 [22] で表される：

$$p(\{r_{t,j}\} | \sigma_r) \propto \exp \left[ -\frac{1}{2\sigma_r^2} \sum_{t,j} (r_{t,j} - r_{t-1,j})^2 \right] \times \exp \left[ -\frac{1}{2\sigma_r^2} \sum_{t,j \sim j'} (r_{t,j} - r_{t,j'})^2 \right] \quad (10)$$

式 (10) の指数関数部分の第二項では、カテゴリー  $j$  についての和は隣接する売場のカテゴリー  $j'$  についてとる。指数関数部分の第一項は時間のつながりを表しており、時間の連続の条件を反映している。指数関数部分の第二項は空間のつながりを表している。ここでは五つのカテゴリーの隣接関係を考慮している (表4)。式 (6)~(10) のシステムは状態空間モデルの行列形式にまとめられ、パラメータの事後分布はMCMCによって推定することができる [23]。本分析ではすべてのパラメータに無情報事前分布を仮定した。推定方法の詳細は文献 [13] をご覧いただきたい。なお、ここで紹介する事例では  $N_{t,j}$  はすべての時点で既知としてモデルに与える。

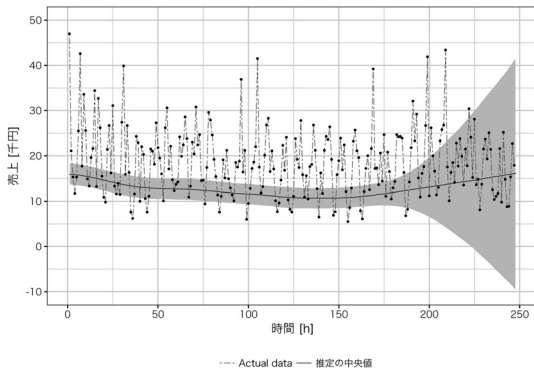
#### 4.2 分析結果

ここでは惣菜エリアの結果のみを示すことにする。図5~図7はモデルによる推定結果を表している。図5は売上のモデルによる予測と実測値を重ね書きしたもので、点が観測値、実線が予測の中央値、帯がベイズ予測区間である。10月の最後の週 (56時間) はテストデータと定め、モデルの訓練には使用していない。図中の丸点が訓練データ、三角印がテストデータ



△ Test data ○ Training data — 推定の中央値

図5 惣菜の10月の売上変動とモデルによる予測



--- Actual data — 推定の中央値

図6 トレンドモデルの推定結果

を表す。ベイズ予測区間は濃い灰帯の50%、薄い灰帯の90%と二種類を図示した。縦軸は千円単位のカテゴリの売上、横軸は1時間単位の経過時間である。推定には9月のデータも含まれるが、ここでは10月の1カ月のデータのみを図示している。日単位(8時間の単位)では周期的な構造が見て取れ、全体的にも緩やかな変化が生じていることがわかるだろう。単純な観測データの可視化だけでは、市場に潜む微細な変動の構造まで見抜くことは難しい。確率分布と組み合わせることで、より鮮明に市場の変動を明らかにすることができる。スーパーマーケットは1日の時間帯のほかにも、休日と平日で売場構造は異なってくると考えられる。このような要因を考慮したい場合は、別々にモデル化し状態空間モデルに組み込む必要があるだろう。

図6はトレンドモデルの推定結果である。縦軸と横軸は図5と同じスケールで、実線が予測の中央値、灰

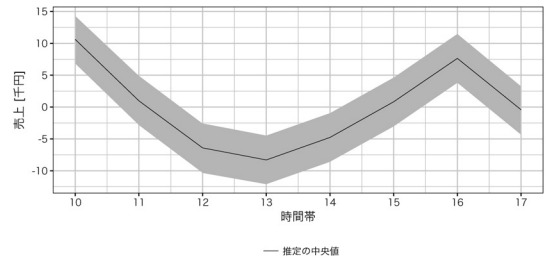


図7 季節調整モデルの推定結果

帯が90%ベイズ信頼区間である。トレンド項は対象とするモデルのスケールにおいて長期的な変動を表すが、月中で売上が落ち込む傾向があることがわかるだろう。図7は季節調整モデルの推定結果で、10月1日を例として載せている。昼の12時~14時は特に落ち込む傾向があることがわかる。図8は売場の時空間構造を示す潜在変数 $r_{t,j}$ の可視化である。縦軸は五つのエリアカテゴリー、横軸は10月の1時間単位の経過時間である。マルコフ確率場で推定されているため、時空間は滑らかな接続となっている。潜在変数 $r_{t,j}$ は式(10)を介して直接的に売上のモデルに組み込まれているため、図8で値の高い場所がそのまま売上への影響が大きいことを示している。時間方向で見ると、明確に周期構造が見て取れる。2(惣菜)と4(一般Aと一般B)で色が明るいことから、これらのカテゴリーは売上への影響が強いといえる。店舗運営の改善策を考える際には、まずはこの二つのカテゴリーに注目するとよいだろう。また、酒、冷食、ドリンクなどのカテゴリーでは訪問数を増やす必要がある。

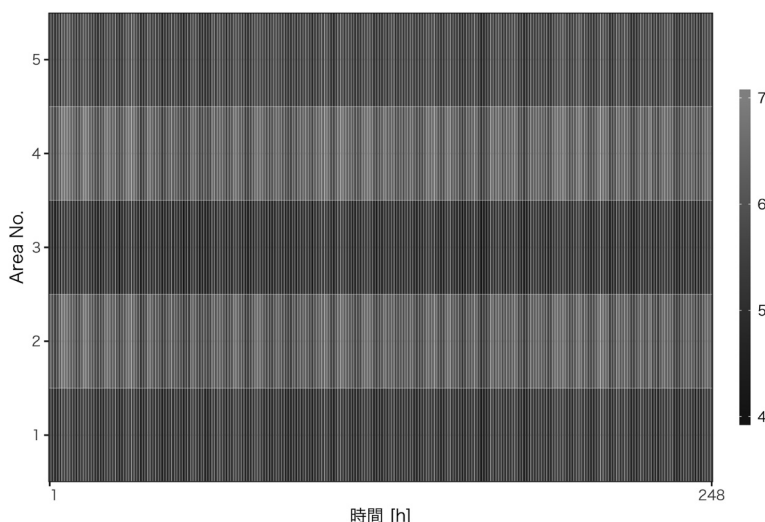


図8 顧客の移動を含む売場の時空間構造の可視化

以上の分析の結果から業務改善で必要となるポイントを列挙することができる。まずトレンドモデルの分析結果から、月中で売上が落ち込む傾向があるため、この近辺では顧客を呼び込む経営上の工夫が必要である。季節調整モデルの分析結果から、昼は売上が落ち込む傾向にある。訪問数も落ちていることから、この時間帯でも顧客数を増やすための経営努力が必要である。酒、冷食、ドリンクの売場は顧客の訪問がほかの売場に比べて少ない。基準の顧客数を増やすための施策が必要と言えらる。

## 5. おわりに

本稿ではスケール階層の視点から、顧客一人のスケールに注目したフラクタルの購買モデルと、集計データのマクロなスケールに注目した状態空間モデルについて解説をした。顧客動線のフラクタル次元を含む2項ロジットモデルで顧客の購買行動を分析した結果、青果、鮮魚、精肉の商品カテゴリーでは、フラクタル次元が高くなるほど顧客は商品を購入しやすくなることがわかった。青果、鮮魚、精肉を購入する主婦層に対しては、商品の陳列や告知を工夫するなどして、売場の回遊を促す経営施策が有効と考えられる。このようにフラクタル次元の導入によって、顧客の回遊の複雑さが購買行動に与える影響を定量的に評価することに成功し、示唆に富む結果を得ることができた。一方、状態空間モデルを用いた分析では、業務改善のために必要な数々の洞察が得られた。トレンドモデルでは売上の長期的な変動を捉え、月中で売上が落ちる傾向にあることを見だし、季節調整モデルでは昼に売上の下

がる時間帯が存在することを明らかにした。そして売場の時空間構造を表す潜在変数を使い、売上への影響の強い売場訪問数を可視化することができた。この可視化された売場の時空間構造のマップは、現場のマーケティングスタッフにも理解が容易であり、経営施策の改善に積極的に利用していくことを提案できる。

さて、ここで実務への実用性から離れ、フラクタル（自己相似性）や状態空間モデルに通底するスケールの概念に立ち戻ってみる。本稿での切り口は、スケール階層の視点から消費者行動を捉えることであった。フラクタルはスケール不変性という性質を有しており、吉田はこれを〈スケールの凝集体〉と表現している [24]。一方で、本稿で紹介した状態空間モデルでは、時間と空間のスケールの連関が式 (10) によって表現されている。マルコフ確率場によって売場の時空間方向の接続が実現されているのである。これは集計データの形で分断されたスケールの構造を回復させる<sup>2</sup>、という意味合いをもつ。「消費者行動の系は非線形性を有する」と考えるならば、集計データから消費者行動を分析することは、“秩序の構造”を抽出しようとする試みであることに對し、非集計データの分析は“無秩序の乱れ”に着目することに相当する。秩序と無秩序の階層が連関し混在すること、それが消費者行動の複雑性の本質である。この複雑性の探求こそ、数多の自然現象にも共通して、われわれの科学が取り組むべき普遍的なテーマであると言えらる。

<sup>2</sup> 自然科学で扱う精緻な微分方程式の数理モデルでは、高階微分に微小係数のかかった“特異摂動項”の存在によってスケールの構造を回復させる。

## 参考文献

- [1] 矢田勝俊, “スーパーマーケットにおける顧客動線分布と文字列解析,” *統計数理*, **56**(2), pp. 199–213, 2008.
- [2] S. K. Hui, P. S. Fader and E. T. Bradlow, “Path data in marketing: An integrative framework and prospectus for model building,” *Marketing Science*, **28**(2), pp. 320–335, 2009.
- [3] S. Miyazaki, T. Washio and K. Yada, “Analysis of residence time in shopping using RFID data: An application of the kernel density estimation to RFID,” In *2011 11th IEEE International Conference on Data Mining Workshops*, pp. 1170–1176, 2011.
- [4] 宮崎慎也, “RFID 技術を用いて取得したスーパーマーケット内の顧客動線の可視化,” *日本建築学会技術報告集*, **18**(40), pp. 1033–1037, 2012.
- [5] Y. Kaneko, S. Miyazaki and K. Yada, “Visualization system for shopping path,” *Procedia Computer Science*, **60**, pp. 1772–1779, 2015.
- [6] B. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, 1982.
- [7] H. Takayasu, *Fractals in the Physical Sciences*, Manchester University Press, 1990.
- [8] 仲佐博裕, 『アコースティック・エミッションの理論と実際—非破壊検査におけるフラクタル評価法—』, 地人書館, 1994.
- [9] Y. Kaneko and K. Yada, “Fractal dimension of shopping path: Influence on purchase behavior in a supermarket,” *Procedia Computer Science*, **96**, pp. 1764–1771, 2016.
- [10] Y. Kaneko, “Fractal analysis of a grocery store shopping path,” In *Proceedings of the 2nd Asia-Pacific World Congress on Computer Science and Engineering*, pp. 1–7, 2015.
- [11] 里村卓也, 『R で学ぶデータサイエンス 13 マーケティングモデル』, 第 2 版, 共立出版, 2015.
- [12] 岩波データサイエンス刊行委員会, 『岩波データサイエンス』, vol. 4, 岩波書店, 2016.
- [13] Y. Kaneko, S. Miyazaki and K. Yada, “The influence of customer movement between sales areas on sales amount: A dynamic Bayesian model of the in-store customer movement and sales relationship,” *Procedia Computer Science*, **112**, pp. 1845–1854, 2017.
- [14] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*, Springer, 1997.
- [15] P. A. Naik, M. K. Mantrala and A. G. Sawyer, “Planning media schedules in the presence of dynamic advertising quality,” *Marketing Science*, **17**(3), pp. 214–235, 1998.
- [16] F. M. Bass, N. Bruce, S. Majumdar and B. P. S. Murthi, “Wearout effects of different advertising themes: A dynamic Bayesian model of the advertising-sales relationship,” *Marketing Science*, **26**(2), pp. 179–195, 2007.
- [17] H. J. Van Heerde, C. F. Mela and P. Manchanda, “The dynamic effect of innovation on market structure,” *Journal of Marketing Research*, **41**(2), pp. 166–183, 2004.
- [18] 山口類, 土屋映子, 樋口知之, “状態空間モデルを用いた飲食店売上の要因分解,” *オペレーションズ・リサーチ: 経営の科学*, **49**(5), pp. 316–324, 2004.
- [19] 北川源四郎, 『時系列解析入門』, 岩波書店, 2005.
- [20] 樋口知之, 『予測にいかす統計モデリングの基本』, 講談社, 2011.
- [21] 本橋永至, 磯崎直樹, 長尾大道, 樋口知之, “状態空間モデルによるインターネット広告のクリック率予測,” *オペレーションズ・リサーチ: 経営の科学*, **57**(10), pp. 574–583, 2012.
- [22] 岩波データサイエンス刊行委員会, 『岩波データサイエンス』, vol. 1, 岩波書店, 2015.
- [23] 松浦健太郎, 『Stan と R でベイズ統計モデリング』, 共立出版, 2016.
- [24] 吉田善章, 『非線形とは何か—複雑系への挑戦—』, 岩波書店, 2010.