

# 顧客の潜在需要とチェーン間競合を考慮した ブランド再配置

経営システム工学専攻 李涵碩

## 1 研究目的

近年、ファッション業界においては、SPA 業態と呼ばれている一つの会社が商品の企画から製造、販売、在庫管理、店舗企画までをすべて一貫に行う企業が成長し、注目を集めている。SPA 業態では、消費者の多様化のニーズを満足するため、多数のブランドを製造し、それを使い分けながら出店する場合が多い。このような企業においては、消費者全体が何を求めているのかを把握するのは当然のこと、各店舗に来店される顧客がどんな潜在的ニーズを持っているかを把握し、それに応じて店舗でのブランド配置が必要であると考えられる。また、チェーン店の場合、特に中心街に出店されている店舗においては、周りの他社ブランドと競争している一方、自社のチェーン店の間にも競合が起きる可能性がある。同じチェーンに属する店舗の間では、競合を避けながらどれだけ顧客を集められるのかが企業全体の売上増加につながる。

本研究の目的は、SPA 業態を採用しているファッションチェーン店の顧客購買データを用いて、顧客の潜在ニーズを把握した上で、より合理的な店舗ブランド選択モデルを提案することで、売上の向上を図ることである。また、本研究が研究対象とする企業は主に中心市街地のショッピングモールや百貨店に出店しているため、一つの商圈に複数の店舗が存在する場合があります、チェーン同士の間にも競合が存在していると考えられる。本研究では、ブランド選択モデルの結果を用いて、ある地域内のチェーン間競合を避けるための施策を提案する。そして、すべての顧客の来店距離での増減を基準とすることで、施策を評価する。

## 2 データ概要

本節では、本研究の分析にあたって使用したデータについて説明する。

本研究の分析対象は SPA 業態を採用しているファッションチェーン店である。この企業は複数のブランドを展開しており、全国のショッピングモ-

ルや百貨店に約 300 店舗を出店している。特徴として、店舗により扱っているブランド・カテゴリやその構成比率が大きく異なることが挙げられる。

本研究では ID-POS データと会員情報データを使用する。用いた項目は以下に示す。

- ID-POS データ  
購買に関するデータ: 売上日, 店舗名称, 顧客コード  
商品に関するデータ: ブランド名, カテゴリ名, 売上数量, 上代売上, レジ売上を使用した
- 会員情報データ  
会員の生年, 顧客番号と登録した店舗

利用できるデータ期間は 2015 年 9 月から 2016 年 8 月までの一年間である。そのうち通年営業をしている店舗のみを分析対象とし、全国に計 299 店舗となる。本研究では、日ごとの掛け率によりデータ期間を 4 期間に分け、そのうちのプロパー期の一つを抽出した。分析対象とするデータは以下の通りである。

- データ期間: 2016 年 3 月 10 日~6 月 14 日
- 購買履歴: 131350 行 (会員データのみ)
- 期間内会員顧客数: 37066 人
- カテゴリ数: 29
- ブランド数: 115
- 対象店舗: 299

また、後の章で説明するように商圈と商圈人口を推定するためには総務省統計局 [2] が提供した以下のデータを用いた。

- 平成 27 年国勢調査 都道府県・市区町村別主要統計表
- 平成 27 年国勢調査 人口等基本集計
- 平成 27 年国勢調査 小地域 (町丁・字等別)

## 3 分析のモデルと手順

本章では、分析のため提案するモデルと研究の手順 (図 1) について述べる。

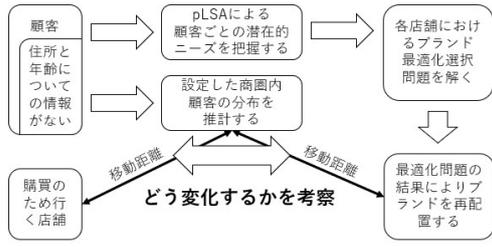


図 1: 分析の流れ

### 3.1 顧客の潜在需要に基づいた店舗のブランド選択モデル

本節では、二項ソフトクラスタリングを用いて、顧客が商品を購入する確率を算出する。顧客を  $x_i$ 、購入したブランド・カテゴリ商品の組合せを  $y_j$ 、顧客と商品が同時に所属するクラスを  $z_k$  とする。この場合顧客  $x_i$  が商品  $y_j$  を購入する確率を

$$p(x_i, y_j) = \sum_k p(z_k) p(x_i | z_k) p(y_j | z_k) \quad (1)$$

で表すことができる [1].

次に、店舗ごとの各カテゴリの潜在需要を算出する。上記の  $p(x_i, y_j)$  を用いることで、店舗  $l$  におけるカテゴリ  $j$  の潜在購買力  $S_{l,j}$  が得られる。

$$S_{l,j} = \sum_{X_i} A_i \times p(X_i, j) \quad (2)$$

ここでは  $X_i$  を店舗  $l$  に来店した顧客、 $A_i$  を顧客  $X_i$  が購入した商品数とする。

上記求めた、各店舗のカテゴリ別潜在的購買力をもとに、全国売上上位の 100 店舗でのブランド選択を行う。ここでは、各ブランドの構成比率を変数とする最小化問題を解く。目的関数は店舗におけるカテゴリごとの潜在的購買数と、ブランドの構成比率でウェイト付けされたブランドごとのカテゴリ別販売数の残差平方和を各ブランドでの総和をとったものである。

最小化

$$\sum_j \left( \frac{S_{l,j}}{\sum_j S_{l,j}} - x_b \times R_{b,j} \right)^2 \quad (3)$$

制約条件

$$\sum_b x_b = 1 \quad (4)$$

$$x_b \geq 0 \quad (5)$$

$$\sum_{b \in B} \geq 0.6 \quad (6)$$

ただし、 $x_b$  はブランド  $b$  の構成比率であり、 $B$  は主要ブランド (売上上位の 11 ブランド) の集合を表す。  $R_{b,j}$  はブランド  $b$  におけるカテゴリ  $j$  の販売数の構成比率であり、これを最小化する。また、上位の 25 ブランドがほとんどの売上を占めているため、ここでは上位 25 ブランドのみを分析対象とした。

### 3.2 顧客年齢と住所の推定

チェーン間の競争を避けるため、各店舗において需要度の高いブランドのみを配置するように提案する。そして、顧客が再配置した前後の移動距離差を基準として、再配置の施策を評価していく。

本データには顧客住所についての情報は含まれていないため、ここで設定した商圏内から一様に来客すると仮定し、商圏を市・区単位に分割し、それらの人口を基準にランダムにどの顧客がどの地域から来たのかを推定した。顧客  $i$  が商圏内の区・市  $k$  から来た確率  $p_k$  は、その区の対象人口数が商圏内の対象人口数と比例している。

$$p_k = \frac{H_{k,a} \times R_k}{H_{total}} \quad (7)$$

ただし、 $H_{k,a}$  は区・市の対象人口数、 $H_{total}$  は商圏内対象人口数の合計値である。この企業は 50 代以上の女性をターゲットとしているため、ここでの対象人口は各区・市に居住している 50 代、60 代と 70 代の日本籍女性とする。また、各区・市がちょうど完全に商圏内に所属するわけではないため、含まれている各市・区の面積がその区・市の総面積を占めている比率  $R_k$  を用いる。商圏に含まれている面積は QGIS を用いて算出した。また、集計結果から区・市により年代別の人口構成比率が異なっていることが分かったため、人口数を年代別で計算する必要がある。

ここで、一部の店舗は規模や立地などの関係で会員数が少ない、また、会員顧客の中でも年齢が適切に記入されていないケースが存在していると考えられるため、非会員も対象として分析に加えた。ただし、非会員についての年齢情報がないため、コレスポネンス分析を用いて購買履歴から年齢層の推測を行った。

本論文では、各行に顧客の年代属性、各列に購入した商品が所属するブランド・カテゴリとした。得られた結果により散布図における購買ごとの各年代属性までの距離  $S_{a,b,age}$  を算出することができる。ここでは  $a$  はブランド、 $b$  はカテゴリを表す。各年代に属する確率比は距離の逆数の比である。

次に、得られた割合になるように一様分布に従う乱数からどの属性が発生させて決定した。そして、前節と同じように、顧客年齢の情報をを用いて、居住地を推測する。

### 3.3 最適化モデルの結果に基づき配置した距離の比較

まずは推定した顧客の居住地を用いて、実際に購買を行った店舗までの距離を計算する。居住地の緯度・経度をその区・市の区役所・市役所の緯度・経度とする。また、すべての顧客の移動距離の総和  $X$  も算出する。

自チェーン内での競争を避けるためには、そのブランドの需要が大きい店舗のみを優先し、扱うようにする。具体的には、最適化モデルの結果に基づき、各店舗におけるブランド構成比率の平均値を求める。そして、平均値より大きい店舗のみを「そのブランドを扱う」、小さい場合は「扱わない」とする。すでに考察した各店舗の行きやすさから、顧客が推定された自宅住所(商圏半径を 5km と仮定する) から求めるブランドを扱っている距離の一番近い店舗に行くことと仮定する(これをケース A と呼ぶ)。次に、すべての購買に対して推定した住所から距離の一番近い店舗までの距離を算出し、距離の総和を  $Y$  とする。

また、比較案として、以下の場合の距離  $X$  と  $Y$  も算出する。

- ケース B: 元々そのブランドを扱っている&最適化の結果において構成比率が平均値以上の場合を「そのブランドを扱う」とする。
- ケース C: 一部の会員顧客のデータのみ&最適化の結果において構成比率が平均値以上の場合を「そのブランドを扱う」とする。
- ケース D: 商圏半径を 10km とした場合&一部の会員顧客のみ&構成比率が平均値以上の場合を「そのブランドを扱う」とする。
- ケース E: 商圏半径を 10km とした場合&一部の会員顧客のみ&元もとそのブランドを扱っている&構成比率が平均値以上の場合を「そのブランドを扱う」とする。

## 4 結果と考察

本節では、提案したモデルと各手順の結果をまとめる。また、結果についての考察を論じる。

### 4.1 結果

図 2 は最小化問題を解いた結果である。横軸には売上の上位 25 ブランド、縦軸には各店舗における構成比率を表す。

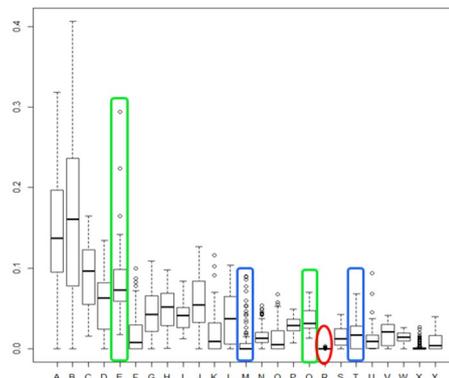


図 2: ブランド再配置の結果

コレスポンデンス分析を行った結果は図 3 に示す。オレンジは購入したブランド、青はスタイル、グレーは顧客の年代属性を表す。表 1 では各ケー

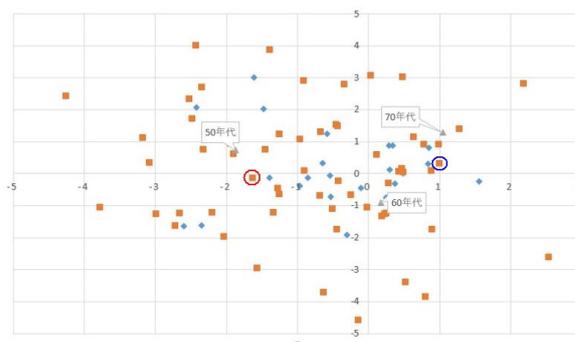


図 3: コレスポンデンス分析の結果

スにおける  $X$  と  $Y$  の値を示す。また、移動距離に変化がある顧客の人数を集計した。

表 1: 実際の移動距離と再配置した後の移動距離

	ケース A	ケース B	ケース C	ケース D	ケース E
元もとの店舗への移動距離総和 $X$	34622872.69	34622872.69	13494280.08	20682791.28	20682791.28
再配置した後の移動距離総和 $Y$	45597690.31	51810218.31	18607356.66	22215437.99	23912452.47
$Y/X$	1.32	1.5	1.38	1.07	1.16
元もとの移動距離平均値	3702.59	3702.59	3749.45	5746.82	5746.82
再配置した後の移動距離平均値	4876.24	5540.61	5170.15	6172.67	6644.20
店までの距離遠くなった人数	2983	3378	1137	789	990
店までの距離が近くなった人数	3545	2872	1335	1325	1112
距離が変わらない人数	2823	3101	1127	1485	1497

### 4.2 考察

最小化問題の解に基づいて、各店舗における顧客潜在的ニーズに応じたブランド再配置を行った。上位ブランドの構成比率が比較的に高く、特にブランド A とブランド B は基幹ブランドであることが分かった。ただし、店舗により構成比率の差も大きくあり、顧客ニーズの多様性を表している。また、全体の売上が低いブランドにしても特定の店舗において重要な役割をに担っていることがある。

表1より、ケースAの場合顧客の総移動距離が実際移動距離の1.32倍となり、顧客ごとの平均移動距離は1173.65mを増えた。図4は顧客の移動距離差の分布を示す。横軸は距離(m)、縦軸は人数を表す。差の分布から移動距離が少し短くなっている顧客が多く、一部の顧客の移動距離が非常に増えているため、全体の距離も大きく増加したことが考えられる。また、各店舗に扱っているブランド数もかなり少なくなったため、顧客ニーズに応じたブランド再配置をすることで、東京都内店舗の競合を避けるながら多くの顧客にもより便利な購買環境を作れたと考えられる。

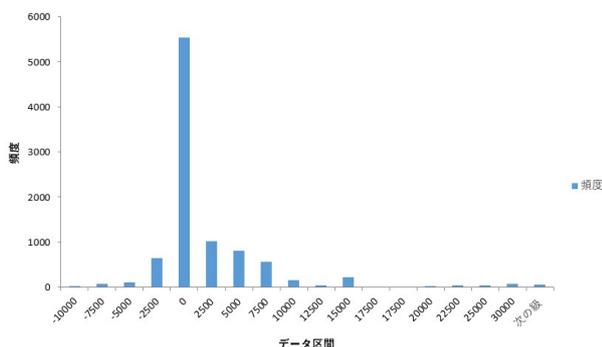


図4: ケースAの場合距離の差  $X - Y$

ケースBの場合顧客の移動距離総和が実際の1.5倍となった。これが制約により再配置した後、一部の店の扱っているブランド数がかなり少なくなったためだと思われる。これらの店舗において、最初扱っていなかった需要度の高いブランドを入れるか、最適化結果により扱っているブランドの中値の高いブランドを残す施策が考えられる。ケースCの場合は、ケースAに年代が50代から70代までの顧客のみという制限がつけられたものである。結果はケースAとほぼ一致していることから年齢の推測はうまくいったと言える。ケースEはケースCの商圈を10kmまで拡大したものである。結果の1.04倍はケースCよりかなり良くなっていった。これにより商圈半径を10kmと設定したほうがより適切であると考えられる。ケースDの得られた距離が実際の0.91倍であった。結果的には一番良いとは言える。

以上のことから、競合を避けるためのブランド再配置は、顧客全体の移動距離差を避けるためであれば、最適化結果により構成比率が平均値以上のブランドのみを扱うとするので良い。店舗間ブランドの入れ替わりの面倒を避けるためであれば、平均値以上かつ実際に扱っているブランドを扱う

とすれば良い。ただし、この場合一部の店舗において扱うブランド数が極端に少ないため、需要度の高いブランドを入れることが必要である。

## 5 結論

### 5.1 本論文のまとめ

本研究は、SPA業態を採用しているファッションチェーン店を対象に、顧客ニーズを把握した上でのチェーン間競合を避ける施策を提案することを目的とした。まずは、実際の購買履歴を用いて顧客の潜在ニーズを算出した。それに基づき各店舗の扱っているブランドの最適化モデルを提案した。また、東京都各店舗の交通利便性から一部の顧客を多店舗に行ってもらうことが可能であるため、需要の高いブランドだけを扱うことでチェーン同士の競合を避ける。そして、実際に購入するまでの距離総和と再配置した後の距離総和を比較することで、再配置の評価を行った。本論文の成果として、顧客の潜在的ニーズを把握することができ、それにより各店舗でのブランドごとの構成比率を最適化することができた。また、チェーン間競合を避けるため、最適化の結果を用いて店舗での扱うブランドを再配置する施策を提案し、顧客の移動距離総和を用いて評価した。

### 5.2 今後の課題

本研究では3月から6月のプロパー期データのみ対象として考えたが、季節により売れてるカテゴリ、ブランドが大きく変わっていく可能性がある一方、扱っているブランドを大きく変えるのは不可能である。そのため、顧客のニーズとブランド選択の間のバランスをどう取るかを考える必要がある。

住所区分を町村まで絞るとより精度の高い結果が出ると考えられる。また、今回は顧客をすべて周辺に居住している場合しか考えていない。アパレル業界のため郊外や遠い地域からショッピングしに来る人も大勢いると考えられる。

## 参考文献

- [1] T. Hofmann, "Probabilistic Latent Semantic Analysis," Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, 289-296, 1999
- [2] 総務省統計局 e-Stat <https://www.e-stat.go.jp/> (最新アクセス日: 2018/02/13)