

実店舗と比較したネットスーパーの併売商品の分析

下津 拓未¹ 吉野 孝¹ 貴志 祥江² 大西 剛² 松山 浩士³
 1 和歌山大学 2 株式会社オークワ 3 株式会社サイバーリンクス

1 はじめに

これまで、ネットスーパーと実店舗の傾向比較は、あまり行われてこなかった。しかし、ネットスーパーの商品の購入発注された際、実店舗の陳列されている売り場から注文された商品を取ってくるため、ネットスーパーの売り上げを考慮せずに店舗のみの在庫だけで管理を行うと、ネットスーパーによって売れた商品の影響により、実店舗では在庫切れといった状態に陥る可能性が考えられる。また、実店舗側でもネットスーパーの大量購買により陳列商品が抜けていると、新規の買い物客から見てもその店舗の商品の品揃えが良くないといった悪い印象を与えてしまい、次の買い物では自店舗では購入してもらえない可能性も考えられる。これらのことを防ぐためにも、実店舗だけでなく、ネットスーパー側の売上や利用ユーザの購買傾向を知る必要があると考えられる。

我々は、ネットスーパー側と実店舗側の比較を行うことで、ユーザ間で購買傾向や特徴に違いがあると仮説を立てた。本研究では、ネットスーパーと実店舗の商品を比較するために対象として POS データの、「客数」「売上金額」「売上数量」「売上種類数量*1」「特売金額」に着目し、ネットスーパーと実店舗の比較を行い、それぞれの傾向の違いについてデータ分析を行う。

2 関連研究

生田目らは、POS データから購買行動に関するデータ分析を行った [1]。この研究では、主に天候の変化を考慮したバスケット分析を行った。本研究とは、バスケット分析を用いてさらに天候を考慮した分析を行っている点が類似している。しかし、POS データを用いたネットスーパーに関する分析は行われていない。本研究では、実店舗だけでなくネットスーパーを考慮した分析を行う。

池田はネットスーパーの特徴分析として、ネットスーパーと実店舗購買率の分析を行った [2]。この研究では、商品を部門ごとで分類している。さらに、ネットスーパーユーザの中から利用頻度ごとに分類を行い、ユーザ間での購買率の分析をしている。本研究とは、ネットスーパーと実店舗の比較分析を商品をもとに調査している点が類似している。しかし、バスケット分析で商品を分類している点で異なる。本研究では、商品を併売関係で分類するため、バスケット分析を行う。

3 POS データの分析

POS データとは、商品がレジで購入される時のデータのことであり、POS データの分析について述べる。

3.1 使用するデータ

使用するデータは、株式会社オークワで収集され、COVID-19 の影響を受けていない 2019 年 8 月 21 日か

ら 2019 年 12 月 20 日の 4 か月分の POS データである。店舗は、和歌山市内にある 1 店舗対象としており、その店舗の POS データを購入方法別に区別するために、データを実店舗での購買データとネットでの購買データの二つに分割した。

POS データ

POS データとは、どの時間帯で購入されたかが分かるデータである。その他のデータとして、日付、RECNO*2、加工コード*3、各商品の部門コードやラインコード、クラスコードなどの詳細情報、JAN コード*4や売上の金額・数量、見切り*5の金額・数量、特売*6の金額・数量、商品の定価が含まれている。

高リフト商品群

商品は期間内でネットスーパーと実店舗の両方で購入された商品のみ限定した。つまり、一方だけで販売されている商品は除いている。それらの商品をネットスーパーと実店舗のデータからバスケット分析を行い、それぞれのデータで併売*7関係が約 500 個ずつの合計約 1000 個の併売関係を抽出した。その中に記されている約 100 点*8の商品をリフト値*9が高い商品を併売傾向が高い商品とみなし、それらの商品の情報が日付ごとに記載されている POS データである。

低リフト商品群

商品は期間内でネットスーパーと実店舗の両方で購入された商品のみ限定した。つまり、一方だけで販売されている商品は除いている。それらの商品の中から、高リフト商品群を除いた商品のみ限定した。商品数でみると約 4,500 点の商品の情報が日付ごとに記載されている POS データである。

3.2 POS データの分析

ネットスーパーと実店舗のそれぞれの高リフト商品群と低リフト商品群から分析した。また、「客数」「売上金額」「売上数量」「売上種類数量」「特売金額」を分析対象とし、それらの要因となっている属性を特定するためにランダムフォレストによる変数重要度分析を行った。変数重要度とは、目的変数を求めるために用いた説明変数の中でそれぞれの変数がランダムフォレストの回帰分類においてどれほど貢献しているかを量的に表した数値である。つまり、変数重要度が高いということは、その目的変数において、高い相関関係があるということであり、ネットスーパーと実店舗の比較においても傾向に違いがあると推測される。

3.2.1 ランダムフォレストによる変数重要度分析

高リフト商品群と低リフト商品群のそれぞれをネットと実店舗に分け、「客数」や「売上金額」だけでなく、「気温」「天気」「曜日」なども説明変数として回帰分析を行った。その際に、どの変数がどれだけ重要視されてい

Analysis of concurrent selling commodities at an online supermarket compared with an actual store

Takumi Shimotsu¹ Takashi Yoshino¹ Sachie Kishi²
 Takeshi Onishi² Koji Matsuyama³

1 Wakayama University

2 Okuwa Co., Ltd.

3 Cyber Links Co., Ltd.

*1 購入された商品数

*2 レシートに 1 枚ずつつけられた番号

*3 RECNO に紐づき、会員である顧客を識別するための番号

*4 商品を識別するためにつけられた番号

*5 通常の価格より値段を下げられてしまった、売れ残り商品

*6 チラシ、タイムバーゲン、朝市、値引き券などで購入された商品

*7 併売とは、2 種類以上の異なる商品が同時購入されることである

*8 併売関係には同じ商品が重複しており、商品の種類は少なかった

*9 ある商品と他の商品の相関を表す指標の一つ

るのかをランダムフォレストの変数重要度でグラフ化を行った。そして、ネットと実店舗の比較、また高リフト商品群と低リフト商品群の比較を行い、それぞれの重要度の変化に着目した。分析の結果、「曜日」で大きな差が見られたので、「曜日」に着目した分析を行った。その際に、図1のような結果が得られた。図1は、「客数」「売上金額」「売上数量」「売上種類数量」それぞれの目的変数に対して、説明変数の中から曜日だけを抽出し、変数重要度を可視化した図である。

図1(1)の「客数」では、水曜日と金曜日において実店舗よりもネットの方が数値が高く、図1(2)では、金曜日と土曜日において実店舗よりもネットの数値が高くなっていることが分かる。これらのことから、「客数」「売上金額」においては、ネットでは水曜日や金曜日などでかなり高い数値が見られ、実店舗ではどの曜日でもほとんど変わらず全体として数値が低いという傾向が見られた。しかし、図1(3)の「売上数量」では、日曜日においてネットよりも実店舗の方が数値が高く、図1(4)の「売上種類数量」では、火曜日や日曜日においてネットよりも実店舗の方が数値が高くなっていることが分かる。これらのことから、「売上数量」「売上種類数量」においては、ネットでは木曜日や金曜日などでとびぬけて高い数値がみられたがそれ以外の曜日ではそれほど高い傾向がなく、実店舗ではネットよりも高い数値がある曜日がみられ、他の曜日でもネットと同じくらい高い数値が表示されていることが分かった。さらに、図1(1)、(2)と図1(3)、(4)でのこれらの数値の違いから「客数」「売上金額」と「売上数量」「売上種類数量」の目的変数によってネットと実店舗の傾向が、違っているということが「曜日」においては判明した。さらに、月曜日や火曜日に比べて、木曜日や金曜日といった週末に近い曜日ほどネットの数値が高い傾向にあることが分かる。これらの結果から、週末であれば、時間的な余裕があるので利用者の都合よい時間帯に自宅で商品を受け取ることが可能といった部分や週末に備えたまとめ買いといったネットスーパー独自の利点が反映されているのではないかと推測される。また、「気温」や「天気」でも同様に分析を行ったが、同様の傾向は見られなかった。これらの結果から、「気温」や「天気」よりも「曜日」にネットと実店舗の潜在的な違いがあると考え、曜日別で客数や売り上げとの深い関係性を見る必要があると考えられる。

3.2.2 曜日別分析

曜日別データを使用し、曜日ごとの割合を線グラフで可視化を行った。高リフト商品群と低リフト商品群で比較し、「客数」「売上金額」「売上数量」「売上種類数量」「特売金額」それぞれを対象に分析を行った。結果より、「特売金額」では、曜日間で大きな差が見られたが、それ以外ではあまり見られなかった。そこで、今回は「特売金額」に着目して分析した。

分析の結果、図2のような結果が得られた。図2は併売ありと併売なしのそれぞれのデータを曜日ごとに分割し、特売金額の合計を一週間を100%とすると、それぞれの曜日が何割を占めているかを数値化し、ネットと実店舗のそれぞれで高リフト商品群と低リフト商品群で表示したものである。図2から併売傾向が高い商品はネットスーパーでも実店舗でも、主に土曜日が高い割合を占めていることが分かった。また、水曜日と木曜日に着目すると、水曜日では、ネットスーパー側では、併売傾向が高い商品の売り上げの割合が増えているが、実店舗側では、併売傾向が高い商品の売り上げの割合が減っていることが分かる。逆に、木曜日では、水曜日とは対照的にネットスーパー側では下がっているが、実店舗側では上がっていることが分かる。このように、曜日によって特売商品の売れ筋が変化するだけでなく、併売関係によっても大きく左右されるということが判明した。このことから、特定の曜日では、特売金額である商品の併売関係の有無により、実店舗とネットでの特売商品を変え

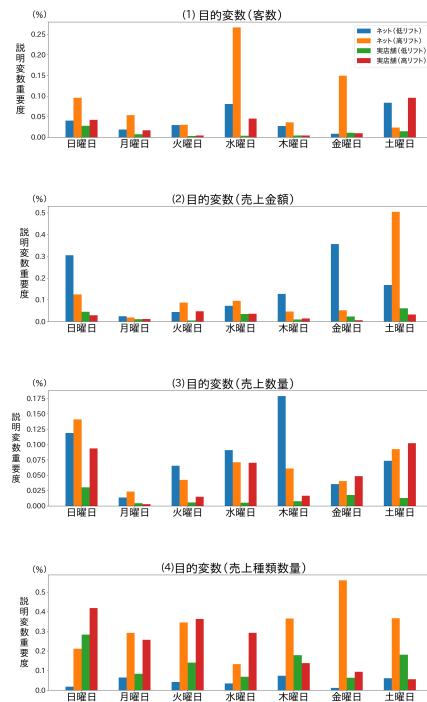


図1 ランダムフォレストによる変数重要度別グラフ

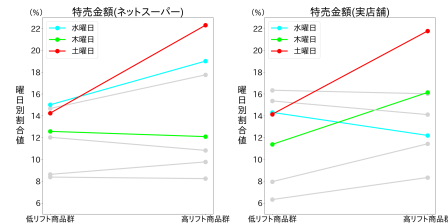


図2 特売金額の曜日別ネットスーパーと実店舗の割合比較グラフ

る必要があると考えられる。また、在庫管理においても特定の曜日では、ネットと実店舗で売れる特売商品は併売関係の有無で変わってくる。よって、それらを考慮したうえでネットスーパー側と実店舗側の双方の在庫管理を行うことで、ネットスーパーの大量買いによる実店舗側の商品不足を防ぐことが可能であると考えられる。

4 おわりに

本稿では、POS データを使用し、ランダムフォレストやバスケット分析を行うことでネットスーパーと実店舗の比較を行った。それぞれ分析として違った結果が見られたので、ネットスーパーと実店舗ではマーケティング戦略も変える必要があることが分かった。今後は、併売傾向から、さらに部門やクラスを考慮した分析を行う。

参考文献

- [1] 生田目崇, 須山憲之: 小売店の購買行動における天気の影響, 専修大学商学研究報, Vol.41, No.8, pp.1-29 (2010).
- [2] 池田満寿次: ネットスーパー利用に関する基本的な特徴と、今後の展望, 流通情報, No.507, pp.47-53 (2014).