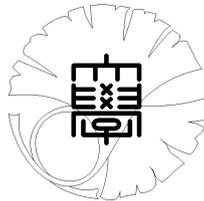


数理科学実践研究レター 2023-8 April 05, 2023

マーケティングにおける統計分析の手法について

by

榎 優一、小池 祐太



UNIVERSITY OF TOKYO

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES

KOMABA, TOKYO, JAPAN

マーケティングにおける統計分析の手法について

榎優一¹ (東京大学大学院理学系研究科物理学専攻)

Yuichi Enoki (Department of Physics, Graduate School of Science, The University of Tokyo)

小池祐太² (東京大学大学院数理科学研究科)

Yuta Koike (Graduate School of Mathematical Sciences, The University of Tokyo)

概要

商品開発にあたって、どのような商品が人気を出すのか、事前に予測することが重要である。本論文では、過去のアンケートデータや売上データに対して典型的な統計手法を適用し、ある程度定量的な形で分析を行う試みについて扱う。

1 はじめに

商品開発に際し、どのような特徴をもつ商品が売れ行きを伸ばすのか、実際に売り出す前のある程度精度良く予測することが望まれる。ここで、予測に関連する客観的で定量的な指標が重要となる。本論文では、そのような指標を得るための試みの一歩として、既存のデータに対して典型的な統計手法を適用して明らかになる性質や適用する際の注意点などについて扱う。

本研究では実際にはエアコンに関するアンケートデータおよび売上データを用いた具体的な分析を行ったが、データ機密の関係上、本論文ではエアコンに限らない商品開発におけるマーケティング分析の手法について一般論を述べることにする。幾分標準的な内容となるが、マーケティング分析の手始めとして、何かの参考になれば幸いである。

2 アンケートデータの解析

購買者の振る舞いを理解するために、しばしばアンケートが実施される。例として、「今回商品を購入した主な理由は何でしょうか？次の選択肢から該当するものを選択してください」という質問事項を考える。様々なメーカーの商品に対して購買者から上記質問事項に対する回答のデータが与えられたとき、主成分分析の手法を用いることで、各メーカーの相対的な立ち位置・特色を捉えることが可能である。具体的には、各会社に対して「各選択肢を選んだ人数」を並べたベクトルが定まるので、それらが属するベクトル空間に座標軸（軸1、軸2、…）を次のように設定し、対応する成分（成分1、成分2、…）ができるだけメーカー間で分散するようにする。

- 軸1は、各メーカーのベクトルの持つ成分1（軸1への射影）の集合の分散が最大となるように設定する。
- 各ベクトルから成分1を取り除いた後、同様に、成分2の集合の分散が最大となるように軸2を取る。
- 軸3以降も同様に取る。

こうして、各メーカーに対して成分1、成分2、…に対する数値が定まる。おおまかには、成分1はそのメーカーの「総合力」を表し、成分2以降は総合力の違いを差し引いた後に現れる、各メーカーの「特色」を表すものとなる。実際にエアコンのデータに対して適用したところ、各メーカーの立ち位置・特色をある程度明らかにすることができた。

¹yuichi.enoki@ipmu.jp

²kyuta@ms.u-tokyo.ac.jp

3 売上データの解析

このセクションでは次の形のデータを扱うことを想定する。

- 商品のリストと機能³のリスト、および各商品が各機能を有しているか否かのデータ
- 一定期間にわたり、各週⁴での各商品の売上台数および売上金額を記載したデータ

商品の売上は季節的な変動を持ちうることに注意しなければならない。そこで上記のデータに対し、二種類の回帰分析を考える。すなわち、季節性を調整した分析（回帰分析 1）と、季節性自体を捉える分析（回帰分析 2）である。

以下では表記の簡単のため、「その商品が機能 i を有するか否か」を表す $\{0, 1\}$ に値を取る変数を χ_i と書く。

3.1 回帰分析 1

エアコンの売上台数や売上金額といった絶対数は、季節によって大きくそのスケールを変化させるため、そのままの値を回帰分析に取り込むと、純粋に機能の有無 χ_i が売上に与える影響を捉えにくい。そこで、回帰分析 1 では、季節性よりも機能の有無 χ_i が売上に与える影響に注目するため、季節性を調整する（季節性の影響をやわらげる）ことを考える。そのために、回帰分析の目的変数を（売上台数や売上金額とはせずに）売上金額の「シェア」とする。具体的には、各週、各商品に対する売上データを用いて、説明変数と目的変数を次のように設定して回帰分析を行う。

- 説明変数：各機能の有無 χ_i 、その週の売上台数 N
- 目的変数：その週の全商品の売上金額の総和に対する、その商品の売上金額のシェア R

すなわち、 a_i, b, R_0 をパラメータとして、以下の式と実際のデータとの差を各データに対して計算し、それらの二乗和が最小になるようにパラメータの値を決定する（最小二乗フィッティング）。

$$R = \sum_i a_i \chi_i + bN + R_0.$$

以下、注意点および考察を述べる。

- 説明変数に機能の有無 χ_i 以外に売上台数を追加することで、機能自体の純粋な影響を抽出することを目指している。今回のエアコンのマーケティング分析では販売価格の差は第一に重要な要素ではないと考えられたため、販売価格は説明変数に追加していない。しかし一般には販売価格の高低が重要となるケースも当然考えられる。その場合、目的変数と同様に、販売価格自体を説明変数に加えると季節変動の影響を大きく受けることがあるため、他の商品との相対的な割合に置き換えて説明変数に加えることが有効と考えられる。
- 機能の有無 χ_i を説明変数とする際、どの機能について考えるべきなのか、データセットに応じた判断が必要である。例えばデータに含まれる商品リスト中で複数の機能が必ずセットで現れる場合は、それらを一つの機能として扱うべきである。また、分析結果の吟味にあたって、各 χ_i に対する回帰係数が単純に「機能 i の有用性」を表していると考えられるべきかどうかとも自明ではない。扱っている商品のリストによっては、 χ_i は実質的には商品の他の特徴・傾向 i' に関する代理変数となっている可能性もある。実際のエアコンの売上データに対する分析でも、このような事例がいくつか見られた。

³ここでは便宜的に「機能」と表記するが、より一般には、商品の特徴を捉える属性を設定してよい。

⁴簡単のため以下で「各週」と表記するが、当然、週毎である必要はない。

3.2 回帰分析 2

次に、季節性を捉えた分析について考える。すなわち、各機能の有無 χ_i に加えて気温や湿度などを説明変数に加え、売上金額自体を目的変数とする回帰分析を考える。説明変数には、各種気候条件の他、すでに考慮した説明変数どうしのクロス項（例えば、気温と機能の有無 χ_i の積）を足すことで、回帰分析の説明能力を上げることができる場合がある。エアコンの売上データに対して実際に行った解析では、気温を説明変数に入れば湿度を追加することによる影響は少なかったが、気温と χ_i のクロス項を追加することには有効性が認められた。具体的には、以下のような回帰モデルを使用した。

$$S = \sum_i A_i \chi_i + BT + CP + \sum_i D_i T \chi_i + S_0.$$

ここで、 A_i, B, C, D_i, S_0 はフィッティングパラメーター、説明変数は各機能の有無 χ_i 、気温 T 、商品の価格 P 、および T と χ_i のクロス項である。

メーカーの選好と機能の選好について エアコンの実際の売上データに対して回帰分析 2 を行った際に観察された事象について述べる。各機能について、気温と売上台数の関係を表す散布図をその機能の有無で分けて表示した際、分布の定性的様子が機能によって大きくふたつのケースに分類された：

- (1) 気温上昇に伴って「機能あり」の商品群のみが売上台数を伸ばし、「機能なし」がほとんど売上台数を伸ばさないケース
- (2) 気温上昇に伴って「機能あり」も「機能なし」も同程度に売上台数を伸ばすケース

これらの違いを検討した結果、(2) のケースに該当する機能の有無はメーカー依存性が高かった。つまり、その機能をいずれの機種にも全くつけていないメーカーもあれば、その機能をほとんどの機種につけているメーカーもあるという状況が観察された。従って、(2) のケースに該当する機能に対する消費者の選好は、メーカーに対する選好を代理していることが予想される。一方、(1) のケースではその傾向は認められず、概ねどのメーカーもその機能がある機種とない機種を共に販売していた。従って、(1) のケースに該当する機能に対する消費者の選好は、純粋に機能に対する選好を捉えていると予想される。

これらの観察から、消費者のメーカーに対する選好は気温の影響を受けにくいのに対し、機能に対する選好は気温の影響を受けやすいという結論が推察される。ここではエアコンの売上データに関する事例を紹介したが、他の商品についても類似の現象が見られる可能性がある。

4 謝辞

本研究に際し、課題を提供していただいたダイキン工業株式会社に感謝する。また、実際に議論に参加して下さったダイキン工業株式会社所属の米田 純也氏、Lal Jaivardhan 氏、東京大学所属の儀我 美一氏、間瀬 崇史氏、吉田 朋広氏および同じグループで様々なアイデアを出してくれた安藤 大悟さん、佐藤 翔一さんにも感謝の意を表す。本研究は数物フロンティア・リーディング大学院 (FMSP) の助成を受けている。