

アイスクリームの販売に関する統計的分析

2000MM051 松下 和哉

指導教員 松田 眞一

1 はじめに

私は、コンビニエンスストアのアルバイトの中で、アイスクリームの発注業務を担当するようになった。発注をする際に、無駄な発注や売り切れを出してしまうことがある。そのため、過去のデータからアイスクリームが売れるための要因を考え、アイスクリームの売り上げ数を予測できないかと考えた。ある程度の売り上げ数を予測することができれば、発注の無駄をなくすことができると考えている。

2 データについて

あるコンビニエンスストアの2002年4月1日から2003年11月31日までの、アイスクリームの販売数を目的変数 y とし、 x_1 = 平均気温・ x_2 = 平均気温の2乗・ x_3 = 日照時間・ x_4 = 降水量・ x_5 = 平均風速 (Web サイト [2] 参照)・ x_6 = J リーグ (Web サイト [4] 参照)・ x_7 = 夏休み・ x_8 = 金曜日・ x_9 = 土曜日・ x_{10} = 日曜日を説明変数としている。アイスクリームに関しては、種類別にカップアイス・コーンモナカアイス・ファンシーアイス・アイスバー・氷菓アイスの5つに分け、個別にエッセル超バナナ・爽バナナ・チョコモナカジャンボ・ピノ・ガリガリ君ソーダ・赤城しぐれいちご・ジャイアントコーン・ジャイアントソフトチョコバナナ・パピコチョココーヒー・パナップの10個を用いた。

3 解析方法

解析方法は、重回帰分析とクラスター分析を用いた。重回帰分析を行うにあたって変数選択法で変数の選択を行った。

3.1 重回帰分析

重回帰分析とは、目的変数 y を p 個の説明変数 x で予測する分析である。

3.2 変数選択法

井上・桑山 [1] の変数増減法と変数減増法のプログラムを利用して変数を選択する。

3.3 クラスター分析

クラスター分析とは、類似度によって似たものに分類する方法である。(田中・脇本 [3] 参照)

4 アイスクリーム全体の分析結果

4.1 変数選択法による変数の選択

変数10個を変数選択法にかけた結果は、自由度調整済み決定係数が0.6803となり、平均気温・平均気温の2乗・日照時間・降水量・Jリーグ・夏休み・金曜日・土曜日・日曜日の平均風速を除いた9個の変数が残った。

4.2 アイスクリーム全体の分析

表1: アイスクリーム全体の分析結果

変数	重回帰係数	標準誤差	t 値	p 値
Intercept	17.60	2.695	6.531	0.0000
x_1	-1.08	0.351	-3.087	0.0021
x_2	0.10	0.012	8.971	0.0000
x_3	0.97	0.170	5.722	0.0000
x_4	-0.15	0.063	-2.418	0.0159
x_6	7.92	4.486	1.767	0.0778
x_7	4.25	2.397	1.772	0.0769
x_8	4.44	1.828	2.427	0.0155
x_9	18.76	1.852	10.126	0.0000
x_{10}	23.18	1.837	12.623	0.0000

アイスクリーム全体の重回帰分析の決定係数は0.69であり、標準誤差は15.09である。重回帰式は、 $y = 17.60 - 1.08x_1 + 0.10x_2 + 0.97x_3 - 0.15x_4 + 7.92x_6 + 4.25x_7 + 4.44x_8 + 18.76x_9 + 23.18x_{10}$ となる。 p 値を見ると、平均気温の2乗・日照時間・土曜日・日曜日がほぼ0となっている。このため、アイスクリームの売り上げに大きな影響を与えると考えることができる。実際に、休日や気温が高い日、日照時間が長い日には店のお客が増えている。そのため、アイスクリームを買っていく人も多くなっていると考えることができる。雨の日には、外を歩く人が少なく、店のお客も少なくなっており、アイスクリームを買って行く人が少ないと考えることができる。

4.3 スチューデント化残差

アイスクリーム全体のスチューデント化残差プロットを図1に示す。

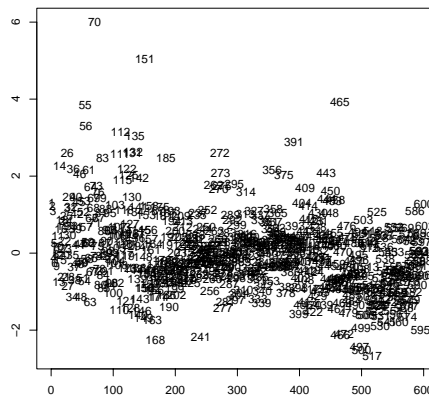


図1: アイスクリーム全体のスチューデント化残差プロット

はずれ値を、大きいものから見ていく。1つ目は2002

年6月9日のワールドカップサッカーの日本対ロシアがトヨタスタジアムで中継されたからだと考えられる。2つ目は、2002年9月1日のモーニング娘。のコンサートがトヨタスタジアムで行われたからだと考えられる。3つ目は、2003年7月12日であるが、大きなイベントがなく集団で買いに来たなどの予測外の出来事が起こったものだと考えられる。

5 種類別アイスクリームの分析

5つの種類別アイスクリームに対して重回帰分析を行った。重回帰式はそれぞれ次のようになった。

● カップアイス $R^2 = 0.33$

$$y = 3.02 + 0.01x_2 + 0.16x_3 + 1.68x_8 + 4.15x_9 + 5.51x_{10}$$

● コーンモナカアイス $R^2 = 0.44$

$$y = 3.61 + 0.02x_2 + 0.17x_3 - 0.04x_4 + 1.93x_6 - 1.22x_7 + 1.45x_8 + 4.81x_9 + 6.06x_{10}$$

● ファンシーアイス $R^2 = 0.35$

$$y = 5.82 + 0.01x_2 + 0.47x_3 - 0.75x_5 + 2.95x_7 + 4.17x_9 + 7.17x_{10}$$

● アイスパー $R^2 = 0.48$

$$y = 4.47 - 0.36x_1 + 0.03x_2 + 0.20x_3 - 0.08x_4 - 0.84x_5 + 2.88x_6 - 2.83x_7 + 0.72x_8 + 4.14x_9 + 3.73x_{10}$$

● 氷菓アイス $R^2 = 0.60$

$$y = 5.36 - 1.20x_1 + 0.05x_2 + 0.14x_3 - 0.03x_4 + 2.38x_6 + 3.10x_7 + 0.93x_8 + 2.17x_9 + 1.08x_{10}$$

各種別アイスクリームの決定係数を見ると、氷菓アイスが0.60となっている他は大きなものとは言えない。種類ごとに一日に売れる個数が少なくなるために、決定係数は低いものになると考えられる。アイスクリームの種類によって残る変数には違いがあるが、すべての種類別アイスクリームで、平均気温の2乗・日照時間・土曜日・日曜日の4つの変数が残った。これらは、アイスクリーム全体でも残った変数であるので、アイスクリームの売り上げを予測するためには不可欠な変数と言うことができる。

6 個別アイスクリームの分析

10個の個別アイスクリームに対して重回帰分析を行った。個別アイスクリームの分析では、決定係数が0.1に満たないものも出てきた。そのため、各個別アイスクリームで残った変数がどのような傾向を示すかを調べた。しかし、スペースの関係により省くものとする。

7 まとめ

種類別アイスクリームを分類するために、アイスクリーム全体の分析で使用した9つの変数で重回帰分析をし、その重回帰係数をクラスター分析にかけた。図2のデンドログラムを見ると、5の氷菓アイスが第1群とすることができ、1のカップアイス、2のコーンモナカアイス、3のファンシーアイス、4のアイスパーを第2群とすることができる。そのため、カップアイ

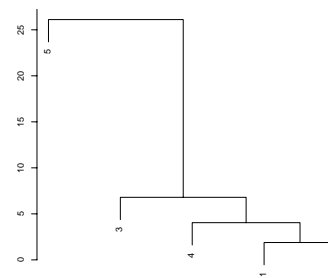


図2: 種類別アイスクリームのデンドログラム

ス・コーンモナカアイス・ファンシーアイス・アイスパーをまとめて重回帰分析にかけて、重回帰式を導き出す。その結果、第1群の重回帰式は氷菓アイスの1種類のみなので、種類別アイスクリームの分析を用い、 $y = 5.36 - 1.20x_1 + 0.05x_2 + 0.14x_3 - 0.03x_4 + 2.38x_6 + 3.10x_7 + 0.93x_8 + 2.17x_9 + 1.08x_{10}$ となる。第2群の重回帰式は、 $y = 15.30 - 0.34x_1 + 0.07x_2 + 0.91x_3 - 0.12x_4 + 6.62x_6 + 3.87x_8 + 17.12x_9 + 22.50x_{10}$ となる。決定係数から、この重回帰分析の説明力は62.34%であるといえる。カップアイスとコーンモナカアイスとファンシーアイスとアイスパーのすべての期間の合計は25851となっており、売り上げ数の比を出すと2.1:2.7:2.9:2.3となる。重回帰式による予測個数をこの比に分けることで各種別の予測個数とできる。この2つの式を発注の参考にするための予測式とする。

アイスクリームは予想通り気温の高さ、日照時間の長さに影響を受けており、平日よりも休日の方が多く売れると言える。Jリーグやイベントがある日には、アイスクリームの売り上げも多くなっているためイベントをしっかりと把握する必要があるとわかった。

8 おわりに

本研究では、アイスクリームの発注を予測する式を提案した。この式を使うことで、今までよりも無駄のない発注に近づけることができるだろう。アルバイトでアイスクリームの発注を行っているので、自分にとって身近なものを研究できて有意義なものであった。

参考文献

- [1] 井上 勤・桑山 智裕, 「S-plusにおける回帰分析の変数選択関数の作成」, 南山大学経営学部情報管理学科卒業論文要旨集, 2001.
- [2] 「気象庁電子閲覧室」, <http://www.data.kishou.go.jp/>.
- [3] 田中 豊・脇本 和昌, 「多変量統計解析法」, 現代数学社, 1983.
- [4] 「トヨタスタジアムでの名古屋グランパスエイト(Jリーグ)の試合」, <http://www.nagoya-grampus-eight.co.jp/>.