

# あるスーパーの部門別売り上げ分析

2003MM065 森山 瑛司

指導教員 松田 真一

## 1 はじめに

私がスーパーで3年近くアルバイトをしている中、売上金額が年々悪くなっているという声をよく耳にするようになった。任される仕事内容も変わり、売上金額を見る機会が増え、売上について考えることが以前よりも多くなった。そこで、この店舗の売上についてもっと知りたい、売上分析をしたいと考えたことが卒業研究のきっかけである。月別や日別の売上金額と日別客数のデータを元にこの店舗の傾向をつかみたいと考えている。また、過去のデータ解析や将来予測を行う方法として時系列解析があるので、並行して学んでいきたい。

## 2 データについて

お店から提供していただくことのできた衣料事業部(5部門)・住居事業部(6部門)・食品事業部(6部門)の2004年7月から2006年8月までの月別売上金額、2005年、2006年の7月12日から2005年10月1日までの日別売上金額と日別客数データとお天気データベース(WeatherEye[2]参照)の2005年、2006年の7月12日から2005年10月1日までの最高・最低気温と天気データを用いる。ただし、住居事業部内のアウトドア・催事・専門店は1部門として考える。

## 3 分析方法

分析方法は、主成分分析、時系列解析(岡田[1]参照)と重回帰分析(船尾[3]参照)を用いる。

## 4 主成分分析

累積寄与率が80%を越す主成分まで取り上げる。

### 4.1 月別売上金額

第一主成分(寄与率0.464)

総合的な売上金額の大きさを表す主成分

第二主成分(寄与率0.159)

相対的に見て衣料品対食料品の売上金額を表す主成分

第三主成分(寄与率0.096)

入学準備や新年度準備の商品対そうでない商品の売上金額を表す主成分

第四主成分(寄与率0.084)

夏対冬の売上金額を表す主成分

#### 4.1.1 考察

月別1日当たりの17部門の部門別売上金額のデータを相関行列を用いて分析した結果、夏や冬といった季節での売上金額の違い、11月から1月の売上金額が伸びる時期や事業部ごとの比較が結果として得られた。店の売れ方として夏や冬、特に冬のボーナス時期には年末商戦として売上金額が一番伸びることがわかる。また、新年度準備品として4月前の3月には、くつ、学生服や化粧品などの商品が売れることがわかる。

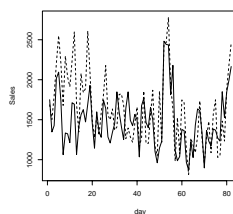
### 4.2 日別売上金額(05年、06年第四主成分までの考察)

2005年、2006年それぞれの日別売上金額データを用いた17部門の部門別売上金額の分析では、女性を表す主成分が多く、来店されるお客さんに女性が多いことがわかる。データの期間が夏ということもあり、お菓子を買うかアイスなどの冷たい商品を買うかという買う商品の違いや食か娯楽かという意識の違いも結果として出ている。また、子供が夏休みということもあり、子供を含んだ主成分が多くみられた。つまり、子供と女性(母親や祖母)という組み合わせで来店されていることや子供を意識した食料品の買い方やゲームやおもちゃなどの文具・玩具の商品の購入がわかる。2005年と2006年では、売れ方に大きな違いはなかった。

## 5 時系列解析(日別売上金額)

全体で売上金額が一番下がっている衣料事業部の日別売上金額を用いる。

### 5.1 時系列グラフと標本自己相関関数



左図をみると2005年と2006年では曜日の違いがあるため売上金額にずれが生じてはいるが、だいたい同じ売れ方をしていることがわかる。衣料事業部の売上金額が2005年に比べて2006年は全体的に下がっている。

図 1: 2005年(点線)、2006年(実線)

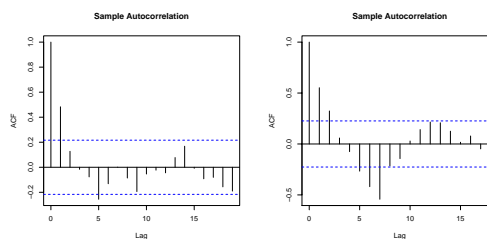


図 2: 標本自己相関(2006年)

図2は、2006年の日別売上金額の標本自己相関の結果(左図)と2006年の日別売上金額から対数をとってラグ7差分を計算した対数差分データ(右図)の結果である。左図は売れ方に周期性が見られないが、右図はキレイな周期性(1週間単位の周期性)が見られる。

### 5.2 モデルへの当てはめ

モデルへの当てはめでは、図2で用いた対数差分データの2006年9月25日から2006年10月1日までのデータを除いた日データを用いる。用いるデータが定常時系列かどうか

かを確かめるためにPP検定を行うと、 $p$ 値が0.01と小さく、単位根があるという帰無仮説は棄却され、定常時系列であることがわかる。

### 5.2.1 ARモデル

日データを用い、Yule-Walker法を用いてARの推定値を求めると、

$$y_t = 0.476y_{t-1} + 0.023y_{t-2} - 0.134y_{t-3} - 0.015y_{t-4} - 0.041y_{t-5} - 0.172y_{t-6} - 0.416y_{t-7} + 0.364y_{t-8} - 0.186y_{t-9} + e_t$$

という関係式が得られた。Ljung-Box検定で残差解析を行うと、 $p$ 値が0.544となり、残差に自己相関はないという帰無仮説は採択される。

### 5.2.2 ARIMAモデル

ARモデルと同じ日データを用いて $p$ 、 $d$ と $q$ を1、 $p$  10、 $0 < d < 1$ と $0 < q < 10$ の範囲で動かしてモデル選択をし、AIC値が最小になるものを選んだ。その結果ARIMA(1,0,7)モデルがAIC=-7.000で選択され、

$$y_t = -0.020 + 0.444y_{t-1} + 0.024e_{t-1} - 0.024e_{t-2} + 0.045e_{t-3} - 0.014e_{t-4} - 0.023e_{t-5} - 0.034e_{t-6} - 0.975e_{t-7} + e_t$$

という関係式が得られた。Ljung-Box検定を行うと、 $p$ 値が0.92となり、残差に自己相関はないという帰無仮説は採択される。

### 5.2.3 考察

モデルへの当てはめでは、次数のみのARモデルよりもARIMAモデルの方が移動平均をとっている分当てはまりがよいといえる。予測では、両モデルとも十分なデータ数とはいえないため正確なものではなかった。

## 6 重回帰分析

### 6.1 日別売上金額

日別売上金額が何に影響しているかを見るために分析を行う。日別客数、最高・最低気温、曜日、天気とダミー変数を用いて分析を行う。また、日別客数を除いた場合も行う。

表 1: 日別売上金額 (客数あり)

項目名	回帰係数	標準誤差	P値
Intercept	0.269	0.059	8.98e-06
客数	0.833	0.033	<2e-16
最高気温	0.031	0.015	0.034
最低気温	-0.020	0.013	0.116
曜日金曜日	0.069	0.026	0.009
曜日月曜日	-0.021	0.027	0.430
曜日水曜日	0.007	0.026	0.787
曜日土曜日	0.191	0.026	9.87e-13
曜日日曜日	0.113	0.026	1.71e-05
曜日木曜日	-0.003	0.026	0.894
天気晴れ	-0.119	0.035	0.001
天気曇り	-0.085	0.032	0.008
年号	-0.043	0.014	0.003
衣料事業部	-0.141	0.078	0.071
住居事業部	-0.596	0.061	<2e-16
決定係数	0.978	調整済み決定係数	0.977

2005年と2006年とで意識改革があったかどうかをみるための説明変数年号が正の値を示しており、接客という意識改革があったといえる。気温では、最高・最低気温が共に1 上がると売上金額は上がる。つまり、暑いと売上金額が伸びることがわかる。曜日では、土日は広告期間中に家族で買い物にきているため売上金額が上がる。事業部では、衣料事業部や住居事業部が食品事業部の売上金額よりも悪いということがわかる。

客数を除いた場合では、食品事業部(食料品)の売上金額が伸び、衣料事業部(衣料品)や住居事業部(住居品)は客数の影響が売上金額に大きく影響することがわかる。

### 6.2 日別客数

日別客数が何に影響しているかを見るために日別売上金額と同様に分析を行う。

表 2: 日別客数

項目名	回帰係数	標準誤差	P値
Intercept	1.400	0.030	<2e-16
最高気温	0.039	0.016	0.015
最低気温	-0.029	0.016	0.066
曜日金曜日	0.057	0.035	0.102
曜日月曜日	-0.131	0.036	0.000
曜日水曜日	0.028	0.035	0.427
曜日土曜日	0.022	0.035	0.538
曜日日曜日	-0.040	0.035	0.255
曜日木曜日	-0.038	0.035	0.284
年号	-0.073	0.019	0.000
衣料事業部	-2.276	0.023	<2e-16
住居事業部	-1.765	0.023	<2e-16
決定係数	0.958	調整済み決定係数	0.957

年号が負の値を示しており、客数の呼び込みが2005年よりも2006年はうまくいかず、客数が減少している。曜日では、日曜日は朝市目当てで求めて来店されるお客さんと考えられ、客数は伸びないと考えられる。事業部では、食品事業部を求めのお客さんが多く、次いで住居事業部と衣料事業部となっている。衣料品目当てで来店されるお客さんが少ないことがわかる。つまり、食料品を中心に買い求めているお客さんが多いことがわかる。

### 6.3 考察

2005年に比べて2006年は客数の呼び込みがうまくいかず、客数の減少が売上金額に大きく影響しているが接客はよくなっている。気温では、暑いと売上金額が伸びることがわかり、2006年は2005年に比べて最高・最低気温の平均が共に1 低く、2005年よりも冷夏であったため売上金額が伸び悩んだと考えられる。曜日では、家族での買い物の休日の売上金額や週末に向かう金曜日に売上金額が伸びる。また、月曜日は土日に必要なものを揃えてしまうため売上金額や客数が伸びない。事業部では、日常で必要とする順、食品事業部、住居事業部、衣料事業部という売上金額の結果となっている。売上金額が下がった原因として、客数の減少と冷夏の影響がわかる。

## 7 おわりに

本研究を通して、店舗の売上金額は全体的に下がっていると予想していたがそうではなく、衣料事業部の売上金額が大幅に下がっていることが店舗の売上金額減少の大きな原因ということがわかった。今後客数確保の対策として、女性と子供を意識したイベントを定期的に行う必要があると考える。時系列解析では、データが少ないためあまり満足のいく結果が得られなかった。

### 参考文献

- [1] 岡田昌史: The R Book データ解析環境Rの活用事例集, 九天社,2004.
- [2] WeatherEye: <http://www.weather-eye.com/>.
- [3] 船尾暢男: The R Tips データ解析環境Rの基本技・グラフィック活用集, 九天社,2005.