

あるホームセンターにおける広告サイズによる売上・客数への影響

2005MM005 坂野智美 2005MM045 森下香緒里

指導教員：田中豊

1 はじめに

本研究では、広告サイズによる売上・客数への影響度を調べることを目的とする。あるホームセンターの現状では、広告サイズが B2 のときでは B3 のときに比べ売上・客数が高くなるということは明らかであった。しかし、B2 では広告経費が高いという問題点があった。そこで、広告サイズによる客数・売上への影響度を統計的に解析し、影響度を正確に把握できれば、広告サイズを変更することで、経費削減ができると考えている。

2 データについて

データは A ~ D 店舗 (A 店舗と B 店舗は名古屋, C 店舗と D 店舗は名古屋以外) の 4 店舗における 2006 年度, 2007 年度 104 週 (各 52 週) の広告サイズと売上・客数のデータである。しかし、週によって広告掲載期間が 4 日間と 5 日間の場合があるので、広告掲載期間をあわせる必要がある。「売上実績 4 日間」、「売上実績 5 日間」、「客数 5 日間」、「跳ね上がり率 4 日間」、「跳ね上がり率 5 日間」、「客数 5 日間 (跳ね上がり率)」という 6 つのデータを作成した。

「売上実績 4 日間」では、広告掲載期間 4 日間および 5 日間のデータについて木曜日から日曜日までの 4 日分を取り出して、売上実績をみている。同様に「売上実績 5 日間」では広告掲載期間が木曜日から月曜日までの 5 日間になっているものだけのデータを用いた。

また各週で広告のある日とない日が存在する。ある日はない日に対してどのくらいの跳ね上がりがあるかを調べることにより季節の影響をなくし、広告の影響を調べていくことを意図としている。ここでは、広告掲載期間である木曜日から日曜日の 1 日当たりの売上実績を、広告掲載期間ではない火曜、水曜の 1 日当たりの売上実績で割ったものを「跳ね上がり率 4 日間」とする。同様に月曜日の売上実績を加えた跳ね上がり率を「跳ね上がり率 5 日間」とする。客数についても同様である。

t 検定では、A 店舗と B 店舗での「売上実績 4 日間」、「売上実績 5 日間」、「跳ね上がり率 4 日間」、「跳ね上がり率 5 日間」のデータより検定を行った。重回帰分析は、A 店舗と B 店舗では上記のデータ全てにおいて分析をし、C 店舗と D 店舗では「売上実績 5 日間」と「客数 5 日間」について分析を行った。

3 分析方法

対応のある t 検定と、重回帰分析の 2 通りで検定と分析を行った。

対応のある t 検定では、広告サイズが B2 の週と B3 の週の差の平均が 0 であるかどうかを検定している。そしてできるだけ他の条件の似た週を対応させ、対応のある 2 標本の差の平均が 0 であれば B2 と B3 の売上、あるい

は跳ね上がり率に差がないことになる。

一方で重回帰分析では、売上実績や跳ね上がり率を説明するために、説明変数に広告サイズを入れ、さらに季節グループ (1 年間を比較的売上が一定と考えられる 8 つの期間, 2 年間では 16 の期間を季節グループとして設定), 広告掲載期間外の火曜水曜の売上を用い、売上や跳ね上がり率にこれらの説明変数がどの程度関係しているかを調べる。また、客数についても同様にして調べる。

4 対応のある t 検定

4.1 対応のある場合の平均の差

広告サイズが B2 の週の売上と、B3 の週の売上の差の平均が 0 かどうかの検定である。

一般に $(X_1, Y_1), \dots, (X_n, Y_n)$ を 2 変量の正規分布からの無作為標本とする。

$$\mu_x = E(X), \quad \mu_y = E(Y)$$

とする。 $\mu = \mu_x - \mu_y$ が推定すべきパラメータである。

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

今回は差が 0 であるかどうか検定したいので、 $\mu_0 = 0$ 。このとき

$$Z_i = X_i - Y_i, \quad i = 1, \dots, n$$

は平均 μ の正規分布からの無作為標本である。この時、帰無仮説のもとで検定統計量は

$$T_0 = \frac{\sqrt{n}(\bar{Z} - \mu_0)}{U_z}$$

を用いればよい。ここで、

$$\bar{Z} = \bar{X} - \bar{Y},$$

$$U_z^2 = \frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^2$$

である。

棄却域 R は

$$R = |T_0| > t_{n-1}(\alpha)$$

とすればよい。ただし $t_{n-1}(\alpha)$ は、自由度 $n-1$ の t 分布の α 点である。

4.2 対応のあるデータの取り方

広告サイズが B2 の週と B3 の週の売上を比較するために t 検定を行う。しかし、時期によって売上の変動があるため、時期の影響がでないように対応のあるデータをとった。各 B3 の週に対して直近の B2 の週を対応させて売上を比較した。すぐ近くの週と対応させることで時期の影響をほとんど受けないだろうという考えである。

4.3 直近のデータの選び方

広告サイズが B3 の週に対して 6 段階の優先順位に従い B2 の週を対応させる。まず第 1 優先として B3 の週の次の週が B2 となっている週を対応させる。次の週が対応さ

せられない場合は前の週を対応させる．それでも対応させられなければさらに2週前後を調べ，最大で3週前後まで調べる．

5 t検定の結果

ここでは，A店舗の「売上実績5日間」と「跳ね上がり率5日間」の結果を示し，他は表にまとめて掲載する．A店舗「売上実績5日間」ではt統計量は2.3006，p値は0.05043となる．また1日当たりの差の平均の点推定量は309,445となる．ここでの1日当たりの差の平均の点推定量は，差の予測値(1,547,224)を広告掲載期間中の日数で割ったものである．

A店舗「跳ね上がり率5日間」ではt統計量は1.3002，p値は0.2297となる．また1日当たりの差の平均の点推定量は451,938となる．ここでの1日当たりの差の平均の点推定量は，広告掲載期間外の売上データの平均に差の予測値(0.1005333)をかけたものである．

5.1 t検定のまとめ

A店舗とB店舗で売上実績と跳ね上がり率のt検定を行った結果をまとめたのが表1である．表1より，跳ね上がり率よりも売上実績の方が信頼できると考えられる．

表1 t検定のまとめ

売上実績	size	1日の差	p値
A店舗4日間	14	280,393	0.1330
A店舗5日間	9	309,445	0.0504
B店舗4日間	13	349,663	0.008064 ***
B店舗5日間	9	274,652	0.03914 *

跳ね上がり率	size	1日の差	p値
A店舗4日間	12	553,018	0.1335
A店舗5日間	9	451,938	0.2297
B店舗4日間	13	410,731	0.05234
B店舗5日間	9	123,555	0.5974

6 重回帰分析

6.1 重回帰分析とは

重回帰分析とは，ある変数 y (従属変数) と，それに影響を及ぼすと考えられる他の変数 x_1, x_2, \dots, x_p (説明変数) に関するデータに基づいて，

$$y_i = a_0 + a_1x_{i1} + \dots + a_px_{ip} + e_i \quad (i = 1, 2, \dots, n)$$

という線形回帰モデルを仮定して，回帰係数 a_0, \dots, a_p を推定する．推定された回帰係数を用いて，説明変数 x から従属変数 y を予測する方法である．

6.2 今回の解析

A店舗とB店舗の「売上実績4日間」，「売上実績5日間」，「跳ね上がり率4日間」，「跳ね上がり率5日間」，「客数5日間」，「客数5日間(跳ね上がり率)」と，C店舗とD店舗の「売上実績5日間」，「客数5日間」を従属変数にして重回帰分析を行った．詳しい結果はA店舗の「売上実績5日間」を掲載し，A店舗の「客数5日間」は簡単に結果を示し，他の結果は表にまとめて掲載する．

6.3 重回帰分析で用いるデータ

6.3.1 従属変数

従属変数には「売上実績4日間」，「売上実績5日間」，「跳ね上がり率4日間」，「跳ね上がり率5日間」，「客数5日間」，「客数5日間(跳ね上がり率)」を用いる．

- 「客数5日間」は，広告掲載期間中(木曜から月曜の5日間)の客数である．
- 「客数5日間(跳ね上がり率)」は，売上実績の跳ね上がり率と同様の計算で，客数を用いて跳ね上がり率を求めたものである．

6.3.2 説明変数

説明変数には，広告サイズ，季節グループ，火曜水曜の売上(客数)を用いる．

- 広告サイズ(B2,B3)は，B2を1，B3を0のダミー変数としている．広告サイズの係数が有意であれば，B2とB3で差があることが分かる．
- 季節グループは，2006年度と2007年度の2年間で16のグループに分けている．季節の影響をなくすために入れている．季節の効果をダミー変数の形で導入することで広告サイズによる売上の違いが分かる．
- 火曜水曜の売上は，広告掲載期間外の売上である．この売上が低いと，跳ね上がり率が高くなると考えられ，この売上が高いと，広告掲載期間中の売上も高いたらうと思われるため，より詳しく従属変数を説明するために用いる．火曜水曜の客数でも同様のことがいえる．

7 A店舗 売上実績5日間

説明変数に広告サイズと季節グループと火曜水曜の売上を用い，それらが売上実績にどれだけ関係があるかを調べるために重回帰モデルを作成し，ステップワイズ法で変数選択を行った結果を表2に示す．

結果

表2 A店舗の売上実績5日間の重回帰分析

	回帰係数	p値
(intercept)	16,120,000	0.000034 ***
広告サイズ	2,530,000	0.000281 ***
季節グループ2	4,886,000	0.012777 *
季節グループ5	-1,035,000	0.190433
季節グループ7	4,619,000	0.002416 **
季節グループ8	-6,112,000	0.001092 **
季節グループ10	2,442,000	0.046156 *
季節グループ15	5,462,000	0.000171 ***
季節グループ16	-6,988,000	0.000108 ***
火曜水曜の売上	0.9384	0.004250 **

広告サイズの回帰係数が2,530,000なので5で割ると，広告サイズがB2のときとB3のときの差額は，1日当たり506,000円となる．

8 t 検定と重回帰分析の比較

A 店舗と B 店舗の「売上実績 4 日間」、「売上実績 5 日間」、「跳ね上がり率 4 日間」、「跳ね上がり率 5 日間」の重回帰分析の結果を表 3 と表 4 にまとめた。

表 3 A 店舗の売上・跳ね上がり率 4,5 日間の重回帰分析

A 店舗	4 日		5 日	
従属変数	size	1 日の差	size	1 日の差
売上実績	74	550,250***	66	506,000***
跳ね上がり率	70	579,575***	67	488,764***

表 1 の t 検定の結果と、表 3 から分かるように、重回帰分析の結果の方が t 検定の結果に比べて差が有意である。これは重回帰分析では広告掲載期間外の火曜水曜の売上が影響しているが、t 検定には広告掲載期間外の火曜水曜の売上が考慮されていないためである。

表 4 B 店舗の売上・跳ね上がり率 4,5 日間の重回帰分析

B 店舗	4 日		5 日	
従属変数	size	1 日の差	size	1 日の差
売上実績	74	395,000***	66	362,200***
跳ね上がり率	74	360,541***	66	304,127**

B 店舗も、表 1 の t 検定の結果と、表 4 から分かるように、A 店舗同様、t 検定に比べて重回帰分析の方が差が有意であることが分かる。t 検定では火曜水曜の売上が考慮されないことも 1 つの要因であると考えられるが、t 検定はサンプルサイズが小さく、データが均一にとれなかったためと考えられる。

これらの t 検定と重回帰分析の結果から、重回帰分析の方が t 検定よりも、より信頼できる分析だと考えられるので、客数の分析は重回帰分析を中心に用いて分析をする。

尚、客数のデータは 5 日間しかないので 4 日間は今後の解析では行わない。

9 A 店舗 客数 5 日間

説明変数に広告サイズと季節グループと火曜水曜の客数を用い、それらが客数にどれだけ関係があるかを調べるために重回帰モデルを作成し、ステップワイズ法で変数選択を行った。

結果

変数選択後に残った変数は、広告サイズ、季節グループ 2,3,5,8,10,14,15,16 と火曜水曜の客数で、広告サイズの回帰係数は 526.7097、p 値は 0.034988* である。

広告サイズの回帰係数が 526.7097 なので 5 で割ると、広告サイズが B2 のときと B3 のときの客数の差は、1 日当たり 105 人 (266,546 円) である。

9.1 客数から売上への換算

あるホームセンターでは、広告による集客力が知りたい。そこで広告サイズが B2 のときと B3 のときの 1 日当たりの客数の差を求めた。そして、客数の差から売上の差へと換算することを考えた。

1 日当たりの、客数の差から売上の差へと変換する方法は、週毎で、1 日当たりの平均売上実績を 1 日当たりの平均客数で割って週毎の客単価を求め、その合計を平均したものを 2 年間の平均の客単価とする。そして、その客単価と 1 日当たりの客数の差を掛けている。

しかし、A 店舗は、売上実績の 1 日当たりの売上の差は 506,000 円で、客数から売上へ換算した 1 日当たりの売上の差は 266,546 円となり、結果が大きく違ってしまった。

ここで、客数と客単価のプロットと、B2 の客単価と B3 の客単価の Box プロットを図 1 に示す。

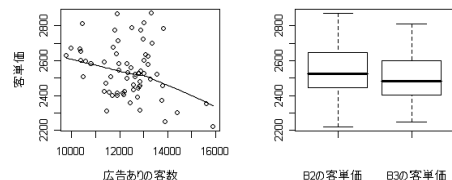


図 1 客数・客単価のプロット

図 1 の左の図より客単価は客数に依存し、右の図より広告サイズが B2 と B3 で客単価に差があるので、一定の客単価を 1 日当たりの客数の差に掛けてはいけなことが分かる。

また、変動係数は、売上実績 5 日間は 1.29 で、客数 5 日間の売上へ換算したものは 2.34 なので、客数から売上へ換算した方がばらつきが大きくなっている。

このことより、B2 と B3 の 1 日当たりの売上の差は、客数からでなく売上実績から求めることとする。

9.2 客数のまとめ

A 店舗と B 店舗の「客数 5 日間」、「客数 5 日間 (跳ね上がり率)」の重回帰分析の結果を表にまとめた。

表 5 A 店舗・B 店舗の客数 5 日間の重回帰分析

5 日間	A		B	
従属変数	size	1 日の差	size	1 日の差
客数	66	105*	66	101**
跳ね上がり率	67	106*	66	76*

表 3 の重回帰分析 5 日間と表 4 の重回帰分析 5 日間と表 5 より、A 店舗に関しては、売上実績や跳ね上がり率の方が、客数や客数 (跳ね上がり率) よりも広告サイズによる影響があると考えられる。

跳ね上がり率は季節の影響を取り除くために用いたが、季節の影響は取り除けなく、自由度調整済み決定係数が全体的に跳ね上がり率よりも実績の方が良いため、C 店舗と D 店舗は実績のみの「売上実績 5 日間」と「客数 5 日間」を用いて重回帰分析を行う。

10 名古屋と名古屋以外の比較

C 店舗と D 店舗の名古屋以外についても、A 店舗と B 店舗の名古屋と同様な結果になるか調べるために、「売上実績 5 日間」と「客数 5 日間」の重回帰分析を行った。ここでは、結果を表にまとめて掲載する。

表 6 C 店舗・D 店舗の売上・客数 5 日間の重回帰分析

5 日間	C		D	
従属変数	size	1 日の差	size	1 日の差
売上実績	67	431,400**	67	369,600***
客数	67	105*	67	114**

表 3～表 6 より、名古屋以外と名古屋で、B2 と B3 の 1 日当たりの売上・客数に大きな差がみられなかった。

11 降水量と最高・最低気温の影響

降水量や気温が変化すると、売上や客数にどのように影響しているか調べるために、売上実績の残差または客数の残差を縦軸に、降水量または最高気温または最低気温を横軸にプロットをし、lowess の平滑化曲線を引いた。

降水量、最高気温、最低気温のデータは、県別の気象庁のデータを用いた。ここでは A 店舗について図を示す。

2006 年度の 15 週と 2007 年度の 19 週の降水量が極端に多く、売上や客数への影響が強いので、この 2 週を外した場合のプロット図を図 2 の右の 2 つに示した。

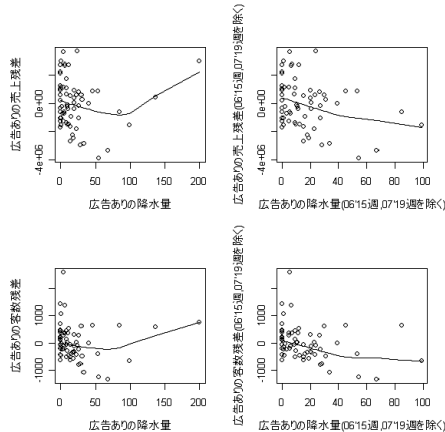


図 2 降水量と売上・客数の残差プロット

降水量が極端に多い 2 週を外した場合の単回帰分析の結果を表 7 に示す。従属変数は売上(客数)の残差で、説明変数は降水量である。

表 7 2 週外した売上・客数の残差と降水量の単回帰分析

	売上残差		客数残差	
	回帰係数	p 値	回帰係数	p 値
intercept	385,986	0.14324	137.762	0.1641
降水量	-26,806	0.00955**	-9.684	0.0126*

降水量に関しては、極端に雨が降った 2 週(降水量 100mm 以上)を除いて、降水量が増すごとに売上・客数は減少すると考えられる。単回帰分析の結果より、A 店舗では降水量が 1mm 増すごとに売上の残差は 26,806 円減り、客数の残差は 9 人少なくなると考えられる。

気温に関しては、最高気温と最低気温では、全体売上・客数に影響がみられないと考えられる。あるホームセンターでは、夏は暑いほど、冬は寒いほど季節の商品がよく売れるという考えがあったが、この結果から、そうで

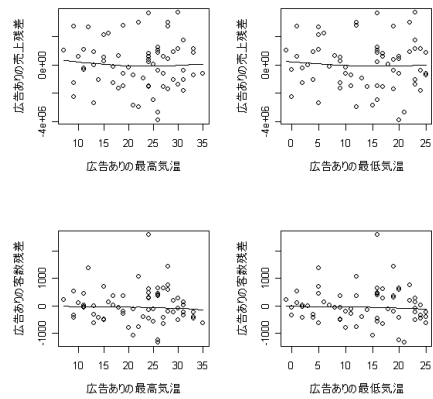


図 3 気温と売上・客数の残差プロット

はないということが分かった。しかしこれは、季節の商品の売上だけでなく、全体の商品の売上で分析をしているからだと考えられる。

12 B2 から B3 へ変更する週の選定

分析を行った B2 の週毎で、売上・客数の残差がマイナスで絶対値が大きい週の広告サイズを B2 から B3 へ変更できると考え、週毎に売上・客数の残差を求めた。降水量が増すと売上や客数が少し減少することも考慮し、4 店舗において 2006 年度も 2007 年度も、11 週、33 週、34 週と 40 週は広告サイズが B2 で残差がマイナスで絶対値が大きいので、これらの週を B3 へ変更できそうだと考えられる。他にも B3 へ変更できそうな週を検討した。

13 おわりに

広告サイズが B2 のときと B3 のときの売上に、差があることが統計的に改めて確認できた。

t 検定ではサンプルサイズが小さく、データが均一に取れなかった。また、広告掲載期間外の火曜水曜の売上が考慮されていないことで十分な結果が得られなかった。

4 店舗で B2 と B3 の 1 日当たりの、売上の差は 304,127 円から 50,6000 円で、客数の差は 76 人から 114 人だった。名古屋と名古屋以外で、B2 と B3 の 1 日当たりの売上・客数に大きな差がみられなかった。

降水量が 1mm 増すと売上は 16,672 円から 26,806 円下がり、客数は 7 人から 9 人減ることが分かった。最高気温と最低気温では全体売上・客数に影響がみられないので、天気は売上や客数にあまり影響がないだろうということが分かった。

4 店舗において 11 週、33 週、34 週と 44 週は B2 の週で、残差がマイナスで絶対値が大きいので、この 4 週は B2 から B3 へ変更できると検討できた。

参考文献

- [1] 金明哲：『R によるデータサイエンス』。森北出版株式会社，東京，2007。
- [2] 白旗慎吾：『統計解析入門』。共立出版株式会社，東京，1992。