

# 顧客ランクによる行動分析

飯塚 久哲, 米村 大介, 豊田 秀樹

## 1. はじめに

スーパーマーケットを中心とする小売店にはさまざまな顧客が存在する。来店頻度という点からみると、毎日来店する顧客もいればたまたま立ち寄っただけの一見の顧客もいる。また、購入商品という点からみても食料品を大量に購入する顧客もいれば、タバコを一箱購入していただくだけの顧客もいる。一口に「顧客」といってもその内容は一様ではなく、さまざまなタイプが存在する。

これら多様な顧客を相手に商業行為を行う際に、すべての顧客に対して同一のアプローチを行うことは顧客満足度の面からみて非効率であり、究極的にはそれぞれの顧客に対して、個別にアプローチをすることが望ましい。そこまで極端でないにせよ、顧客をいくつかの観点から同質のニーズや特徴を持つ、いくつかの集団に分けて顧客を捉えた方がより効率的に顧客の満足度を高めることができる[1]。

そこで、本稿では「月の総購入金額」という観点から顧客をランク分けし、ランクごとに顧客の購買行動を捉えることにより、顧客についての理解を深め、今後の販売活動に寄与することのできる知見を導くことを目的とする。

## 2. 顧客ランク

### 2.1 顧客ランクの定義

まず、本研究における顧客分類基準である「顧客ランク」について説明する。

スーパーマーケットにおける「よい顧客」という条件には多様な見方があるが、経営側からみると購入金

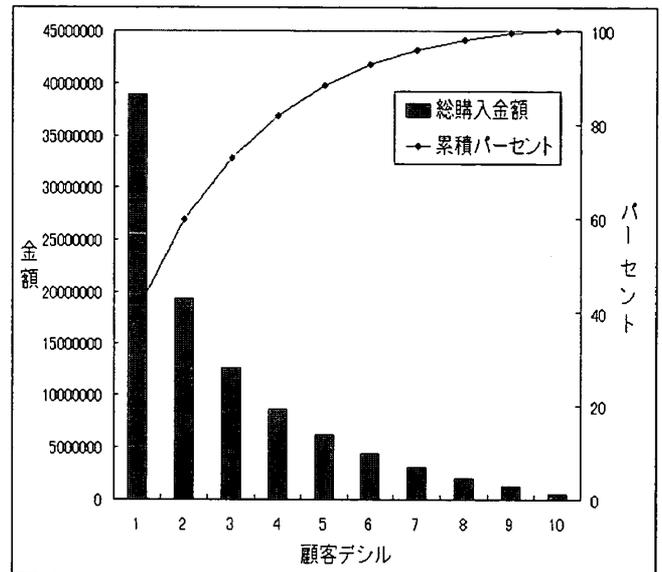


図1 総購入金額のデシル分析結果 (5月)

額の高い顧客、つまり店舗の売上高に貢献する顧客がよい顧客である。そこで、各顧客の1ヶ月間の購買金額がどのような性質を持っているかを検討するために、5月における各顧客の総購入金額に対してデシル分析<sup>1</sup>を行ったところ、図1のようになった。図1をみると、上位30%が総売上高の70%強を占め、中位30%が20%弱を占め、下位の40%が10%弱を占めていることが分かった。そこで、上位30%を「ランク3」、次の30%を「ランク2」、残りの40%を「ランク1」と顧客ランクを定義した。

### 2.2 潜在混合分布モデルによるクラスタリング

前節で定義された顧客ランクは、デシル分析により顧客を三つのランクにクラスタリングするものであった。デシルをまとめて3ランクに分類した理由は、後の分析でランクによる購買行動の差異を比較検討する際に、比較する集団の数が多すぎるとモデル構成がき

<sup>1</sup> デシルとはそれぞれ10%ずつに分割されたグループのことであり[2]、全顧客を購入金額の高い順に10等分してその構成比を算出し、売上の上位集中度の観点から、顧客の全体概況を把握する手法。

いいつか ひさてつ、よねむら だいすけ

早稲田大学 文学研究科

〒162-8644 新宿区戸山1-24-1

とよだ ひでき

早稲田大学 文学部

〒162-8644 新宿区戸山1-24-1

受付 02.8.12 採択 02.11.20

わめて困難になることと、ランクが多すぎると得られた知見を実践に移すのが困難であると考えられたからである。しかしそれだけでは主観的な側面が強いので顧客の5月の総購入金額と総来店回数の2変数による潜在混合分布モデル<sup>2</sup>によって、クラスタリングを試みた。その際、解を得やすくするために総購入金額の値は1000分の1倍して平均と分散を小さくした。分析にはMplus<sup>3</sup>を使用した。

分析の結果は表1の通りである。表内のCは全クラス数を、cはCにおけるクラス番号を表している。また、 $\bar{x}_i$ は平均を、 $\sigma_i^2$ 、 $\sigma_{ij}$ はそれぞれ分散・共分散を、 $r_{ij}$ は相関を表している。3クラスまで解を求めることができたが、4クラス以上では解を得ることができなかった。AIC (Akaike's information criterion)<sup>4</sup>を考察すると、クラス数が3のときが、もっともデータとの適合がよいことが分かる。しかし、表2に示されているデシルによるクラスタリングによって

得られた値とはクラスごとの平均・分散などの推定値が異なっているので、各クラスを構成するオブザベーションは2種類のクラスタリング方法でまったく同じというわけではない。

潜在混合分布モデルによる3クラスはデシル分析の結果をもとに定義した3クラスとは平均・分散・共分散などの推定値が異なるので同一のクラスタリングができたとはいえないが、この結果から、顧客を1母集団として捉えるよりは、優良・非優良の2集団として分類したほうが、2集団として捉えるよりは、非優良をさらに二つに分けて計3集団に分類したほうが、データとの適合がよいという傍証が得られ、デシル分析による3クラスタリングが妥当であったといえる。

### 2.3 顧客ランクの性質

続いて、デシル分析によって定義した「顧客ランク」の性質について調べるために、総来店回数と年齢を外生変数、総購入金額を内生変数とした重回帰分析をランクごとの同時解析で分析した。分析にはAMOS 4.0<sup>5</sup>を使用した。得られた各パラメータ値は表2のとおりである。 $x_1$ が総来店回数、 $x_2$ が年齢、 $x_3$ が総購入金額である。 $\alpha_{ij}$ は変数jから変数iへの回帰係数を表している。このモデルは飽和モデルなので、GFIなどの適合度指標は算出できない。表2から、ランクが上がることにより3変数の平均・分散( $\sigma^2$ は総購入金額の誤差分散を表す)が大きくなること、各ランクとも総来店回数と年齢の間の相関は小さいことが分かる。また、標準化係数 $\alpha_{31}^*$ をみると、総来店回数と総購入金額との関係は、ランク3>ランク1>

表1 潜在混合分布モデルによるクラスタリング

	C=1		C=2	
	c=1	c=1	c=1	c=2
AIC	122550		104088	
人数	9289	5570	3719	
構成比	1.00	0.60	0.40	
$\bar{x}_1$	10.46	3.54	20.83	
$\bar{x}_2$	5.08	2.01	9.67	
$\sigma_1^2$	172.08	7.18	239.91	
$\sigma_2^2$	29.65	1.43	36.77	
$\sigma_{12}$	57.12	2.24	59.96	
$r_{12}$	0.80	0.70	0.64	
C=3				
	c=1	c=2	c=3	
AIC	99139			
人数	3181	3760	2348	
構成比	0.34	0.41	0.25	
$\bar{x}_1$	1.87	7.53	26.79	
$\bar{x}_2$	1.27	3.85	12.19	
$\sigma_1^2$	1.39	16.19	272.38	
$\sigma_2^2$	0.20	3.64	38.47	
$\sigma_{12}$	0.22	3.95	53.02	
$r_{12}$	0.42	0.51	0.52	

( $x_1$ : 総購入金額,  $x_2$ : 総来店回数)

表2 各ランクごとの総来店回数, 年齢, 総購入金額の関係 (重回帰モデル)

	ランク 1	ランク 2	ランク 3
$\sigma_{12}$	1.40	6.39	9.70
$r_{12}$	0.10	0.18	0.12
$\alpha_{31}$	479.37	318.24	1442.35
$\alpha_{31}^*$	0.43	0.40	0.60
$\alpha_{32}$	-1.80	-7.55	-94.76
$\alpha_{32}^*$	-0.03	-0.05	-0.08
$\bar{x}_1$	1.57	3.96	10.87
$\bar{x}_2$	49.47	50.80	52.14
$\bar{x}_3$	1219.22	6052.06	14697.79
$\sigma_1^2$	0.90	6.23	40.14
$\sigma_2^2$	213.39	198.21	174.35
$\sigma_3^2$	891110.13	3406020.60	150760190

( $x_1$ : 総来店回数,  $x_2$ : 年齢,  $x_3$ : 総購入金額)

<sup>2</sup> 連続的な観測変数が予測変数であり、質的な潜在変数が基準変数であるモデル。基準変数そのものが不明な判別分析といえる[4]。

<sup>3</sup> 構造方程式モデルのソフトウェア。詳しくは<http://www.statmodel.com>を参照。

<sup>4</sup> モデル選択の際に利用される指標であり、値が小さいほどよいモデルであると判断する。

<sup>5</sup> 構造方程式モデルのソフトウェア。詳しくは<http://www.smallwaters.com/amos>を参照。

ランク2の順に大きいことがみてとれる。

以上の結果から、顧客ランクの性質を「来ないし買わないランク1」、「来ても買わないランク2」、「来たら買うランク3」というように表現することができる。

### 3. 顧客ランクによる行動分析

顧客ランクが定義できたところで、具体的に顧客ランクという観点から顧客の行動を分析する。その際、(1)来店曜日、(2)購買内容という二つの観点から顧客の購買行動を捉える。来店曜日と購買内容の2点に着目した理由は、小売店における顧客の購買行動として我々が興味を持つのは、「いつ、誰が、何を購入したのか」であり、ここで、「誰」の部分をもつ三つの顧客ランクに固定した際、「いつ」と「何を」がどのように変化するかを把握することができれば、顧客の購買行動の特徴を把握することができると考えたからである。

#### 3.1 顧客ランクと曜日との関係

まず、顧客はランクにより来店する曜日に違いがあるかどうかについて分析した。もし来店曜日が異なるのであれば、店舗では曜日ごとに顧客への対応を変える必要がある。また、来店日が異なるのであれば今後の分析にも曜日の影響を考慮しなければならない。そこで5月におけるランクごとに各曜日の来店回数の合計を算出し、その店舗へ来店した顧客の全来店回数における各曜日の比率を求めた(図2)。図2をみると、ランクによらず火曜日がもっとも顧客の来店回数が多く、次に多いのが月曜日、もっとも来店回数が少ない

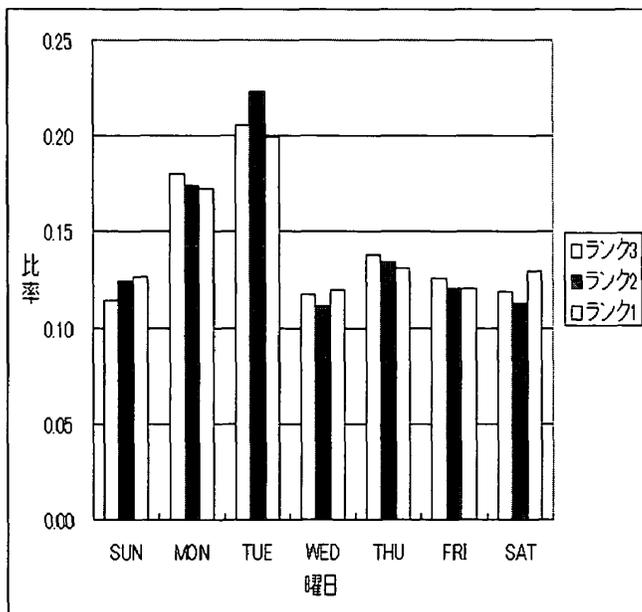


図2 顧客ランクと曜日別の来店回数比率(5月)

のは水曜日であった。ランクによらず、来店する曜日に違いはなく、すべてのランクにおいて、特売日である火曜日への来店が多いことが分かった。

ランクによって来店する曜日に差がないことを確認するため、5月中に来店した曜日を表す七つのカテゴリカル変数を作成して、これらを独立変数とし、顧客ランクを従属変数とした決定木分析を行った。分析にはClementine 6.0.2<sup>6</sup>を使用した。最初は剪定度<sup>7</sup>を60%で行ったが、ルールは導かれなかった。結局、剪定度をどんなに低くしてもまったくルールは導かれなかった。この結果からも、ランクによって来店する曜日に違いがないことを支持する結論を得た。

#### 3.2 顧客ランクによる購買行動の差異

続いて2番目の分析として、ランクによる顧客の購買行動の違いについての分析を行った。つまり、ランクの高い顧客は食料品を購入し、ランクの低い顧客はギフト品を購入するというような傾向があるかどうかについて検討した。

その際、顧客の購買行動を1ヶ月単位で捉えることとし、各顧客の5月1ヶ月間の全商品ラインの購買回数(顧客の各来店日の各商品ラインの購買の有無を顧客に関して総和したもの)を算出し、各ランクごとに積率<sup>8</sup>行列を求め、多母集団の同時解析により、ランクごとの購買行動の違いについて検討した。まず、予備解析として各ランクごとに探索的因子分析<sup>9</sup>を行った。その結果、すべてのランクにおいて、固有値の大きさ、スクリー基準<sup>10</sup>などにより、4因子が妥当であることが示唆された。そこで、各ランクとも4因子で、かつすべてのランクにおいて因子を測定する観測変数が同じになるように変数選択を行った。その結果、図3に示すようなモデルに到達した。

f<sub>1</sub>はミズモノニツパイ<sup>11</sup>、ヤサイ、センギョ、ネリモノニツパイ<sup>12</sup>、エンカン<sup>13</sup>という、食事の際に食卓を飾るおかずを作るときに材料として用いられる商品(観測変数)によって測定されているので「食事因子」

<sup>6</sup> データマイニングの統合的ソフトウェア。

<sup>7</sup> 生成された決定木の剪定の精度を決定するパラメタ。値を大きくするとより小さく簡潔なツリーが得られ、値を小さくするとより精度の高いツリーが生成される。

<sup>8</sup> 変数*i*変数*j*の積率  $m_{ij} = \sum_{k=1}^N x_{ik}x_{jk}$  で表される。平均偏差化されていないため、共分散よりも情報が多い。

<sup>9</sup> 因子負荷がすべて自由母数である因子分析モデル。

<sup>10</sup> 固有値の落ち込みの大きさを決定する方法。

<sup>11</sup> 漬物、豆腐、油揚げなど。

<sup>12</sup> うどん、ソーセージ、ちくわなど。

<sup>13</sup> 珍味、海草、魚のひらきなど。

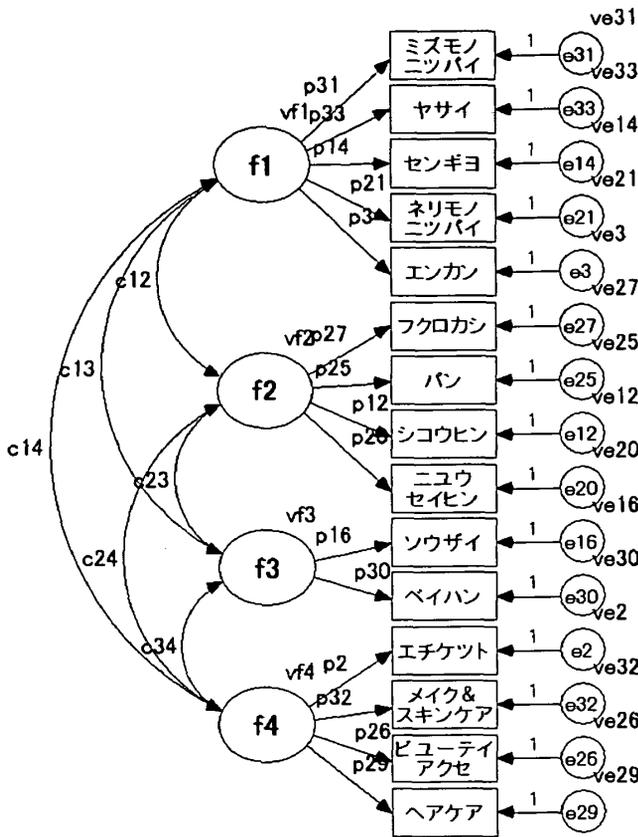


図3 確認的因子分析モデル

と解釈した。

$f_2$ はフクロガシ、パン、シコウヒン<sup>14</sup>、ニユウセイヒンという、 $f_1$ のように食事として利用されるのではなく、おやつとして利用される商品（観測変数）によって測定されているので「おやつ因子」と解釈した。

$f_3$ はソウザイ、ベイハンという、出来合いのご飯（のおかず）によって測定されているので「惣菜因子」と解釈した。

$f_4$ はエチケツト<sup>15</sup>、メイク&スキンケア<sup>16</sup>、ビューティアクセ<sup>17</sup>、ヘアケアという、 $f_1$ から $f_3$ までの食品に関するものではなく、日常生活において必要な生活雑貨品によって測定されているので「生活用品因子」と解釈した。

続いて、モデル1「配置不変モデル=3群で因子を測定する観測変数が等しい（背後に仮定される因子の数や意味が3群とも等しい）モデル」、モデル2「弱測定不変モデル=3群でパス係数（因子負荷）の値が等しい（仮定される因子から観測変数への影響が3群とも等しい）モデル」、モデル3「測定不変+因子の

<sup>14</sup> お茶、即席麺、フルーツ缶詰、レギュラーコーヒーなど。

<sup>15</sup> 日焼け止め、髭剃りなど。

<sup>16</sup> 化粧品、フェイスクケア用品、メイクアップ用品など。

<sup>17</sup> ヘアアクセサリ、ルーム小物など。

表3 モデル選択

モデル	モデル 1	モデル 2	モデル 3	モデル 4	モデル 5
$\chi^2$	2096	2542	16187	70672	106937
df	252	274	294	304	324
GFI	0.970	0.964	0.695	0.528	0.435
AGFI	0.957	0.952	0.626	0.441	0.372
CFI	0.917	0.898	0.284	0.000	0.000
RMSEA	0.028	0.030	0.076	0.158	0.188
AIC	2312	2713	16318	70784	107009

表4 3母集団の確認的因子分析

	ランク 1	ランク 2	ランク 3
$r_{12}$	0.52	0.54	0.80
$r_{23}$	0.19	0.28	0.56
$r_{34}$	0.09	0.09	0.21
$r_{13}$	0.04	0.03	0.39
$r_{24}$	-0.03	0.08	0.48
$r_{14}$	-0.03	-0.02	0.44
$\bar{f}_1$	-1.16	0.00	4.19
$\bar{f}_2$	-0.74	0.00	2.36
$\bar{f}_3$	-0.45	0.00	1.49
$\bar{f}_4$	-0.07	0.00	0.26
$\sigma_1^2$	0.12	0.73	12.97
$\sigma_2^2$	0.05	0.33	4.36
$\sigma_3^2$	0.11	0.59	5.14
$\sigma_4^2$	0.01	0.02	0.14

CFI=0.971, RMSEA=0.03

分散共分散が等しい（因子・因子間のばらつきの程度が3群で等しい）モデル」、モデル4「強測定不変モデル=測定不変+誤差分散が等しい（3群間で等しい因子が3群すべてにおいて同様に測定される）モデル」からモデル5「すべてのパラメタが等しい（3群の分散・共分散行列が等しい）モデル」というように、徐々にモデルに強い制約を課すことにより、各ランクごとの共分散構造の等質性を検討し、モデル選択を行った（表3）。AICをみるとモデル1「配置不変モデル」がもっともデータとの適合がよい。しかし、他の適合度指標も参照すると、モデル2「弱測定不変モデル」まではデータとの適合は保たれ、それ以上強い制約を課すとモデルとデータとの適合が急激に悪くなる。そこで、モデル2「弱測定不変モデル」を採用し、モデル2に平均構造を導入したモデルを最終モデルとして採択した。モデル識別のため、ランク2の因子平均を0に固定することにより、それとの比較でランク1とランク3の因子平均を推定させた。最終モデルのパラメタおよび適合度指標は表4のとおりである。なお、ここではランクによる購買行動の差異に興味があるので、表4にはパラメタとして因子の平均・分散・共分散を掲載し、測定方程式部分のパラメタ（パス係数と

誤差分散)は掲載しなかった。パス係数はすべて有意であり、誤差分散にマイナスの値を示すものはなかった。観測変数は因子を適切に測定しており、かつ、不適解は生じず、正常に解は収束した。

構造方程式モデルによる分析によって見出されたランクごとの購買行動の違いを考察する。まず、ランクが上がることによって、各因子の平均・分散が大きくなっている。つまり、ランクが上がるにつれ、「 $f_1$ : 食事因子」, 「 $f_2$ : おやつ因子」, 「 $f_3$ : 惣菜因子」, 「 $f_4$ : 生活品因子」ともに、たくさん買われる傾向にあり、さらに、人によるばらつきも大きくなる。次に購買行動の違いとして挙げられるのは、因子間相関の大きさの違いである。ランク1・ランク2では「 $f_1$ : 食事因子」と「 $f_2$ : おやつ因子」の間の相関は高いが、他の因子間相関はいずれも総じて低い。しかし、ランク3になると、すべての因子間相関において高い値を示している。つまり、1ヶ月という単位で購買行動を捉えたとき、ランクによらず、ミズモノニツパイを購入する人は、ヤサイやセンギョ、ネリモノニツパイ、エンカンといった食卓品を購入する傾向にあり、フロガシを購入する人は、パン、シコウヒン、ニユウセイヒンといったお菓子類を購入する傾向にある。また、ソウザイを購入する人は、ベイハンも購入する傾向にあり、エチケツトを購入する人は、メイク&スキンケア、ビューティアクセ、ヘアケアといった生活用品も購入するというように、同一の因子内での同時購買は行われるが、ランク1・2の顧客は、因子を超えての同時購買、例えば、ミズモノニツパイとヘアケアといった商品同士の同時購買は行われない。しかし、ランク3では、因子内での同時購買のみならず、因子をまたいでの同時購買、例えば、ヤサイとヘアケアといった商品の購買行動が行われるということである。つまり、「 $f_1$ : 食卓因子」, 「 $f_2$ : おやつ因子」, 「 $f_3$ : 惣菜因子」, 「 $f_4$ : 生活品因子」といった因子ごとに購買行動をするか、あるいは因子を超えて、よりバラエティーに富んだ買い物をするかの相違である。

#### 4. マーケティングアクションの提案

本分析により得られた知見をもとに、実際場面で役立つようなマーケティングアクションを提案する。

本研究で得られた知見は大きく二つ挙げることができる。まず一つめは、「顧客ランクにより来店する曜日に違いはない」ということであり、二つめは、「顧客ランクによる購買行動の差異とは、平均構造という

観点では、ランクが上がるにつれ各因子内での購買行動が活発になること、共分散構造という観点では、ランクが上がるにつれ購買行動のばらつきが大きくなること、そして、ランクが上がるにつれ因子内のみならず、因子のくくりを超えた商品同士の同時購買行動がみられるようになる」ということである。

つまり、優良顧客(たくさん来店し、たくさん購入する顧客)と、非優良顧客(来店も購入もしない顧客・来店しても購入しない顧客)との大きな違いは、購入する商品のバラエティーの広さにあるということである。この点に注目して表4をみると、 $f_2$ と $f_3$ の間の相関係数 $r_{23}$ の値が非常に興味深い。他の因子間相関がおおよそ「ランク2→ランク3」の間で値が高くなっているのとは異なり、 $r_{23}$ は「ランク1→ランク2→ランク3」とランクの上昇とともに着実に相関が高くなっている。

ここから、 $f_1$ 「食事因子」の製品と $f_2$ 「おやつ因子」の製品は、ランクが1から2へと上がり総購入金額が増加しても同時に購買される傾向は強くないが、 $f_2$ 「おやつ因子」の製品と $f_3$ 「惣菜因子」の製品はランク1ではそれほど同時に購入されないが、ランク2ではそれらの同時購買傾向が高くなっている。総購入金額の増加に占める同時購買傾向の変化量が他の組み合わせと比較して大きい、もっとも同時購買されやすい製品カテゴリー群であるといえる。

図4~6はそれぞれランク1からランク3の $f_2$ 「食事因子」と $f_3$ 「おやつ因子」の因子得点をSASのCALISプロシジャ<sup>18</sup>で算出し[3]、各ランクごとにプロットさせたものである。図4から図6をみると、ランクが上がるにつれて直線的関係が強くなっていることがみてとれる。そこで、「おやつ因子」内の製品と「惣菜因子」内の製品を同時に購入してもらえるようなアクションをクーポン券やチラシを利用して行えば、潜在的な顧客の需要を満たすことができ、総購入金額の少ないランク1・ランク2の顧客の購入商品のバラエティーの拡大につながり、各顧客の購入量および購入金額を増加させることが可能となる。その際、来店する曜日についてはランクによる違いがないので、曜日ごとにターゲットとする顧客ランク用に棚割や商品配列を変更するといった特別な配慮をする必要はない。

私たち人間は食事を摂らなければ生活していけないし、生活していく以上身だしなみなどにも注意する必

<sup>18</sup> 構造方程式モデルのソフトウェア。文献[5]が詳しい。

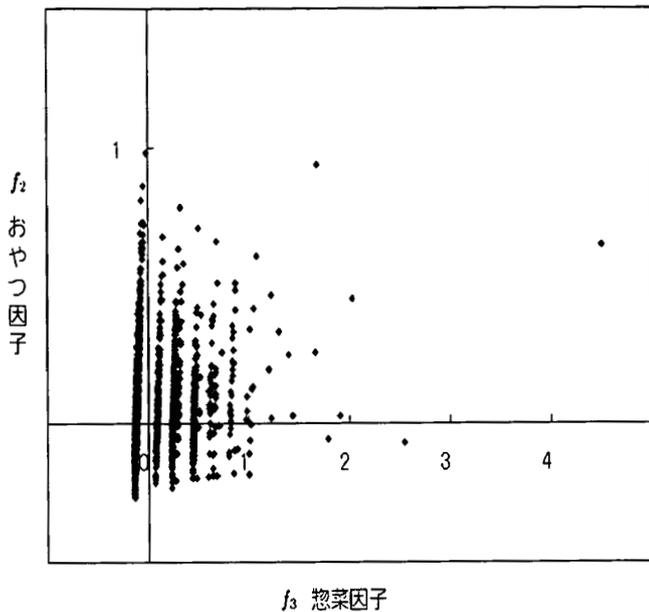


図4 ランク1における $f_2$ と $f_3$ の散布図

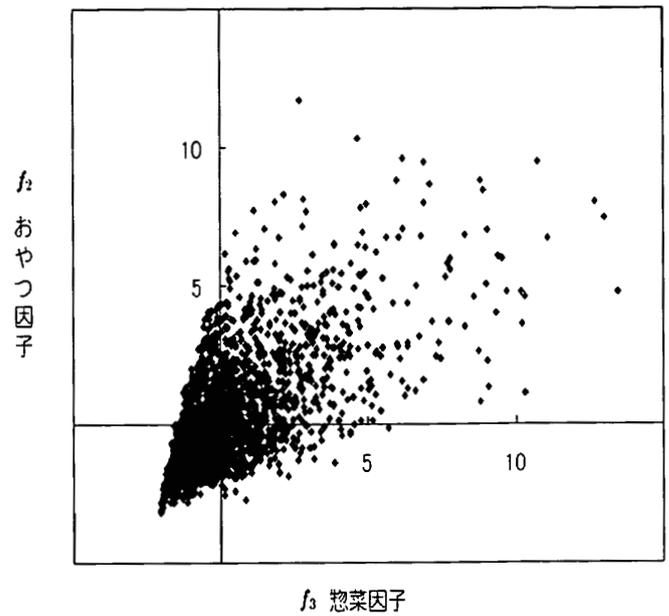


図6 ランク3における $f_2$ と $f_3$ の散布図

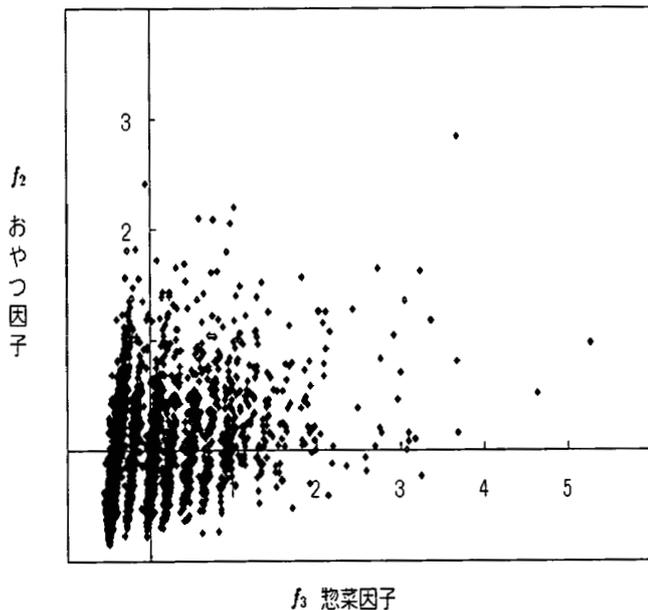


図5 ランク2における $f_2$ と $f_3$ の散布図

要がある。したがって、食品や生活雑貨などは必ず購入しなければならないし、何らかの商品を購入しに来店しているのであるから、この顧客との関係をより良好なものにできれば、顧客のメインスーパーのスイッチングを引き起こすことができ、その結果として、店舗の売上の増加につなげることが可能となろう。

## 5. おわりに

本稿については、立教大学の守口剛先生をはじめとするCRM研究会の皆様、岡太彬訓先生をはじめとする日本OR学会マーケティング・エンジニアリング研究部会の皆様、および、NTTデータ・技術開発本部の中川慶一郎様から多くの助言をいただきました。ここに感謝の意を表します。

### 参考文献

- [1] 朝野熙彦・上田隆穂(著)：“マーケティング&リサーチ通論”，講談社サイエンティフィック，(2000)。
- [2] Blattberg, R. C., Getz, G. & Thomas, J. (著)小川孔輔・小野譲司(監訳)：“顧客資産のマネジメント—カスタマー・エクイティの構築—”，ダイヤモンド社(2002)。
- [3] 鈴木督久・長田公平(著)豊田秀樹(編)：“共分散構造分析 [事例編]”，北大路書房，(1998)，pp. 22-33。
- [4] 豊田秀樹(著)：“共分散構造分析 [応用編]”，朝倉書店，(2000)。
- [5] 豊田秀樹(著)竹内啓(監修)：“SASによる共分散構造分析”，東京大学出版会，(1992)。