

金融チャネル利用実態からの顧客セグメンテーション

里村 卓也, 江原 淳, 佐藤 栄作, 佐藤 忠彦, 寺田 英治

1. はじめに

個別顧客へ対応したマーケティングがさまざまな業界で注目されている。これは銀行といえども例外ではなく、顧客に適した商品を提供（マーケティング）するために顧客DBの活用方法について検討してみよう。ここでは顧客DBとして金融チャネル利用データを活用することを考える。

チャネルとは財およびサービスの提供経路である。銀行の一般利用者にとってはATM（Automated-Teller Machine）は銀行と接することが最も多いチャネルであろう。ある程度自由な時間に、入出金・振込みを行える利便性は非常に大きい。最近では、電話やインターネットを利用した金融サービスの登場で、ATM以外にも銀行顧客が利用可能なチャネル数は拡大しつつある。さらに、またわが国においても流通業の銀行業務参入も始まり、流通業者が母体となる銀行では小売店舗へATMが設置され、預け入れ払い戻し業務に焦点を絞った運営形態がとられている。ATMの設置は金融機関にとって投資がかさむが、利用されやすいATMを持つ銀行にとっては他金融機関からの手数料を得ることができるため、どのようなATM網を持つのかという点での競争力も注目されている。

では、銀行を利用する顧客はどのようなチャネル利用を行っているのだろうか。また利用実態から個別顧客へ対応したマーケティング施策を検討することは

できないだろうか。この問いに答えるために、本研究では金融チャネルの利用実態を利用した顧客のセグメンテーションとマーケティング施策についての検討を行った。

個別顧客に対応したマーケティングの実践のためには個別顧客毎の特性を把握する必要がある。しかし個別顧客への提供物の設計はある程度類型化される必要がある。それゆえ個別顧客対応のマーケティングといえども、顧客をセグメント化して理解することが必要であるといえる。

本研究ではセグメンテーションを行うベースとして金融チャネルの利用実態を用いる。金融チャネルの利用実態を利用することにより識別可能性、到達可能性[2]等、マーケティング戦略での実行を考慮したセグメンテーションを行うことが可能となる。

以下では顧客の月間チャネル利用頻度をモデル化し、潜在クラス分析を適用することにより顧客のセグメンテーションを行い、金融チャネルの利用実態についての考察を行っている。

2. 金融チャネル利用実態からのセグメンテーション

2.1 顧客行動にもとづくセグメンテーション

本研究では顧客の銀行取引データという顧客行動のデータを利用する。顧客別取引データは銀行が自らの全顧客について持っている膨大なデータである。このようなデータは行動データであるので、測定収集が容易である一方、顧客が他の金融機関に持っている資産や金融商品に関する選好は分からない。

そこで、金融チャネル利用実態とアンケートデータを組み合わせることにより、顧客の金融チャネル利用データから顧客のストック（金融商品）選好を予測することを考える。つまり自動的に収集されるフローのトランザクション（金融チャネル利用）から、本来は分からないはずのストックの選好を推定しようというものである。

さとむら たくや

大阪大学 大学院経済学研究科
〒560-0043 大阪府豊中市待兼山町1-7

えはら あつし

専修大学 ネットワーク情報学部
〒214-0033 川崎市多摩区東三田2-1-1

さとう えいさく, さとう ただひこ, てらだ えいじ
財団法人 流通経済研究所
〒141-0031 品川区西五反田7-23-1

2.2 潜在クラス分析

顧客の行動にもとづくセグメンテーションを統計的に行うために、潜在クラス分析を行う。潜在クラス分析では、顧客は各セグメントに確率的に所属すると考える。セグメントが持つ反応や行動特性は特定の関数に支配されていると考える。

本研究では顧客の入出金行動をもとにセグメンテーションを行う。入出金行動はチャネル利用実態である。セグメントが持つ行動特性は各チャネル利用の頻度であると考えられる。ところで、ATM利用頻度等の事象の発生回数は毎月変動するので、データは確率変数の実現値として扱う必要がある。

3. 顧客セグメンテーションのための分析モデル

3.1 利用頻度のモデル化

顧客 $i (=1, \dots, I)$ が単位期間 $t (=1, \dots, T)$ にチャネル $j (=1, \dots, J)$ を利用をする頻度 y_{ijt} はパラメータ φ_j のポアソン分布に従うと仮定する。利用頻度が y_{ijt} である確率は

$$\Pr(y_{ijt}|\varphi_j) = \frac{\exp(-\varphi_j)\varphi_j^{y_{ijt}}}{y_{ijt}!} \quad (1)$$

となる。顧客 i 内で各チャネルの利用頻度は独立であると仮定し、 $\mathbf{y}_{it} = (y_{i1t}, \dots, y_{iJt})$ 、 $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_J)$ とすると、利用頻度が \mathbf{y}_{it} である確率は

$$\Pr(\mathbf{y}_{it}|\boldsymbol{\varphi}) = \prod_{j=1}^J \frac{\exp(-\varphi_j)\varphi_j^{y_{ijt}}}{y_{ijt}!} \quad (2)$$

となる。ところで単位時間で頻度がポアソン分布に従うということは入出金間隔の無記憶性を仮定していることとなる。しかしATMへの振込み、給与振込み、自動引き落とし等月単位の周期性がある可能性がある。もし、利用の発生が周期的であるのであれば出金間隔がポアソン分布であると考えすることは非常に厳しい仮定である。ただし今回の分析で用いたデータは半年毎に1ヶ月間の利用頻度を収集したものである。従ってデータ期間が連続していない。また異なる利用費目や利用目的を全て同一に扱って集計されたデータであるため周期性の影響は無視しうると判断した。

3.2 潜在クラスモデル

顧客によりポアソン分布のパラメータ φ_j は異なるはずである。そこでセグメント $s (=1, \dots, S)$ に所属する顧客 i の持つポアソン分布のパラメータを φ_{sj} とする。セグメント s に所属する顧客 i の利用頻度が \mathbf{y}_{it} である確率は

$$\Pr(\mathbf{y}_{it}|\boldsymbol{\varphi}_s, i \in S) = \prod_{j=1}^J \frac{\exp(-\varphi_{sj})\varphi_{sj}^{y_{ijt}}}{y_{ijt}!} \quad (3)$$

となる。ただし $\boldsymbol{\varphi}_s = (\varphi_{s1}, \dots, \varphi_{sJ})$ とする。ここで顧客は各セグメントに確率的に所属するものとする。顧客 i は確率 π_s でセグメント s に所属するとすると利用頻度が \mathbf{y}_{it} である確率は

$$\Pr(\mathbf{y}_{it}|\boldsymbol{\varphi}) = \sum_{s=1}^S \pi_s \Pr(\mathbf{y}_{it}|\boldsymbol{\varphi}_s, i \in S) \quad (4)$$

である。ただし $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_S)$ と定義しなおす。

$\mathbf{y}_i = (\mathbf{y}_{i1}, \dots, \mathbf{y}_{iT})$ 、 $\boldsymbol{\pi} = (\pi_1, \dots, \pi_S)$ とし、 $t = (1, \dots, T)$ の間に顧客 i のセグメント s への所属確率が変化しないとすると利用頻度が \mathbf{y}_i である確率は

$$\Pr(\mathbf{y}_i|\boldsymbol{\varphi}, \boldsymbol{\pi}) = \sum_{s=1}^S \pi_s \left[\prod_{t=1}^T \Pr(\mathbf{y}_{it}|\boldsymbol{\varphi}_s, i \in S) \right] \quad (5)$$

となる。

全顧客についての期間中のチャネル別利用頻度 $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_I)$ が観測された場合の尤度は次のようになる。

$$L = \prod_{i=1}^I \Pr(\mathbf{y}_i|\boldsymbol{\varphi}, \boldsymbol{\pi}) \quad (6)$$

3.3 パラメータの推定

最尤法を用いてパラメータの推定を行う場合、(6)式には潜在パラメータである $\boldsymbol{\pi}$ が含まれているため、対数尤度が複雑になりニュートン・ラフソン法などの最適化手法では、数値的に安定した解が得られない。そこで本研究では潜在クラス・モデルの推定で一般的なEMアルゴリズム[1]を用いて、パラメータの推定を行った。

EMアルゴリズムでは観測された不完全データをいったん、最尤方程式に馴染みの良い「完全データ」に擬似的に置き換え (E-step)、この「擬似的完全データ」を用いてパラメータの擬似最尤推定値を求め (M-step)、さらに得られた推定値から擬似的完全データを作り直し、それを用いてまたパラメータ推定値を求め直す、という手続きを繰り返す。不完全データに対してEMアルゴリズムを用いることにより得られた推定値は最尤推定値である。

EMアルゴリズムを用いるために、消費者 i がセグメント s に所属するか否かを示す変数 z_{is} を考えよう。消費者 i がセグメント s に所属する場合には $z_{is} = 1$ であり、それ以外の場合には $z_{is} = 0$ である。 z_{is} は観測されない潜在変数である。またセグメント数は所与とする。

このときの完全対数尤度は

$$\log L_c = \sum_i \sum_s z_{is} \log \pi_s$$

表1 各セグメントのセグメント・サイズとチャンネル利用頻度

セグメント名	セグメント サイズ (π_s)	月平均利用回数				
		ATM 平日 時間内	ATM 平日 時間外	ATM 休日 時間外	窓口	センター
フロー・ゼロ	32.4%	0.02	0.00	0.00	0.09	0.13
センター・ヘビー	21.4%	1.63	0.01	0.05	0.37	6.97
ライト・ライト	13.4%	0.83	0.00	0.02	0.17	0.95
ATM ゼロ	13.2%	0.09	0.01	0.00	0.75	3.96
ヘビー・ヘビー	7.9%	5.42	0.09	0.36	0.29	9.10
ATM メイン	7.6%	2.92	0.06	0.16	0.19	1.72
ATM 時間外ヘビー	4.1%	2.55	1.12	1.15	0.17	5.55

$$+\sum_i \sum_s z_{is} \log \Pr(\mathbf{y}_i | \boldsymbol{\varphi}_s, i \in s) \quad (7)$$

ただし

$$\Pr(\mathbf{y}_i | \boldsymbol{\varphi}_s, i \in s) = \prod_{t=1}^T \Pr(\mathbf{y}_{it} | \boldsymbol{\varphi}_s, i \in s) \quad (8)$$

ここで観測されない潜在変数 $Z = ((z_{is}))$ を欠損データとみなすことにより EM アルゴリズムを適用する。

E-step: E-step ではパラメータ $\boldsymbol{\varphi}, \boldsymbol{\pi}$ を固定して潜在変数 Z に関する $\log L_c$ の期待値を求める。

$$E_z[\log L_c] = \sum_i \sum_s E[z_{is} | \mathbf{y}_i] \log \pi_s + \sum_i \sum_s E[z_{is} | \mathbf{y}_i] \log \Pr(\mathbf{y}_i | \boldsymbol{\varphi}_s, i \in s) \quad (9)$$

ただし

$$E[z_{is} | \mathbf{y}_i] = \frac{\pi_s \Pr(\mathbf{y}_i | \boldsymbol{\varphi}_s, i \in s)}{\sum_s \pi_s \Pr(\mathbf{y}_i | \boldsymbol{\varphi}_s, i \in s)} \quad (10)$$

である。このように $E[z_{is} | \mathbf{y}_i]$ は π_s を事前確率としたときの事後確率として求められる。

M-step: M-step では $E_z[\log L_c]$ を最大化する $\boldsymbol{\varphi}, \boldsymbol{\pi}$ を求める。ただし

$$1 \geq \pi_s \geq 0, \quad \sum_{s=1}^S \pi_s = 1, \quad \varphi_{sj} \geq 0 \quad (11)$$

とする。ラグランジュの未定乗数法を用いることにより $E_z[\log L_c]$ を最大化する $\boldsymbol{\varphi}, \boldsymbol{\pi}$ が求まる。

4. 金融チャンネル利用データの分析

4.1 データ

以下のデータを用いて分析を行った。

●銀行取引データ

- 当該銀行での利用チャンネル別の普通預金入出金データ (月次)
- 属性データ (年齢)

●アンケートデータ

- 保有金融資産量

— メインバンク選択理由

— 定期預金について

— 興味のある金融商品について

— 貯蓄の運用を見直すタイミング

●分析対象チャンネル

— ATM 平日時間内利用頻度¹ (回/月)

— ATM 平日時間外利用頻度 (回/月)

— ATM 休日時間外利用頻度 (回/月)

— 窓口利用頻度 (回/月)

— センター利用頻度² (回/月)

●分析対象期間

— 1999年3月, 1999年9月, 2000年3月の各1ヶ月間

●分析対象者

— サンプル数: 3014名 (全サンプル)

— 利用が0であるという情報も重要であるため, 分析対象期間のチャンネル利用の有無により分析対象を分けなかった。

4.2 パラメータの推定とセグメント数の決定

セグメント数をアприオリに与え EM アルゴリズムによりパラメータを推定した。情報量基準の値とマネジリアルな視点から検討した結果, セグメント数を7とした。

4.3 推定結果

セグメント別のセグメント・サイズとチャンネル別の平均利用頻度は表1のとおりである。各セグメントの特徴をまとめた。

- フロー・ゼロ: 入出金等の利用がないセグメントである。サンプル中で最も人数が多い。

¹ 利用頻度は入金と出金の合計

² 給与・年金の振込みや公共料金の自動引き落としなど, 顧客が直接操作しないもの。

- センター・ヘビー：センターの利用が多いセグメントである。ATMの平日時間内利用も月に平均1.6回ある。
- ライト・ライト：フローの利用が皆無ではないが非常に少ないセグメントである。
- ATMゼロ：ATMの利用はほとんどなく、センターの利用が月に平均4回弱あるセグメントである。
- ヘビー・ヘビー：ATM時間内利用、センター利用の利用頻度が最も多いセグメントである。
- ATMメイン：ATMの利用頻度がセンターの利用頻度よりも多いセグメントである。
- ATM時間外ヘビー：平日および休日のATM時間外利用頻度が最も多いセグメントである。

4.4 セグメント別の顧客の特徴

チャンネル利用頻度以外のセグメントの顧客の特徴をまとめる。

- ATM利用1回あたりの利用金額³：「ATM時間外ヘビー」は1回あたりの利用金額が少ない。時間外利用金額は約2万円である。「ATM時間外ヘビー」はATMを財布代わりに使っている可能性が大きいといえよう。「センター・ヘビー」「ヘビー・ヘビー」「ATMヘビー」とも1利用あたりの時間内利用金額に大きな差はない(図1)。
- 当該銀行での資産保有状況：「センター・ヘビー」は当該銀行で資産保有量が最も多い。「フロー・ゼロ」は当該銀行で資産保有量が最も少ない(図2)。
- メインバンク選択理由：セグメント別のメインバ

³ 利用金額は入金と出金の合計

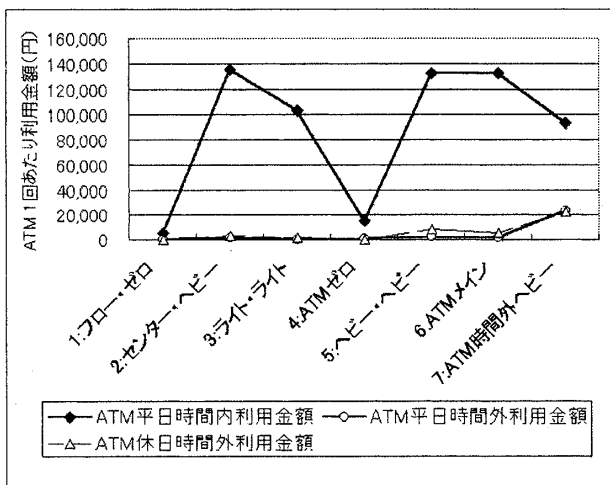


図1 ATM利用1回あたりの利用金額

ンクの選択理由を図3に示す。「給与振込み口座がある」はセンター利用回数が多いセグメントに多い。また「長年取引がある」は「ATM時間外

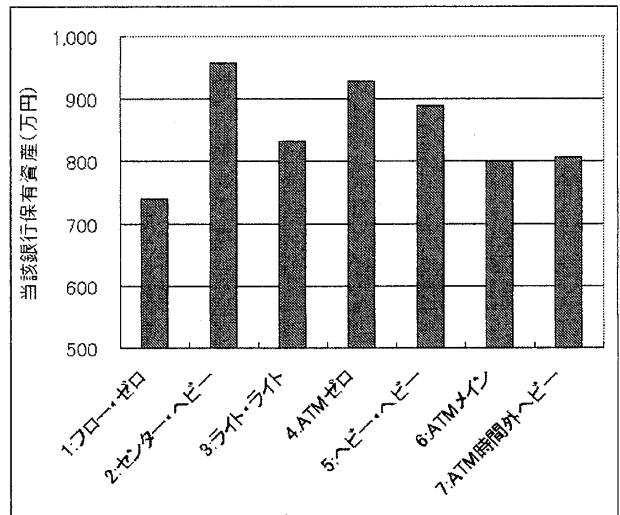


図2 当該銀行での資産保有状況

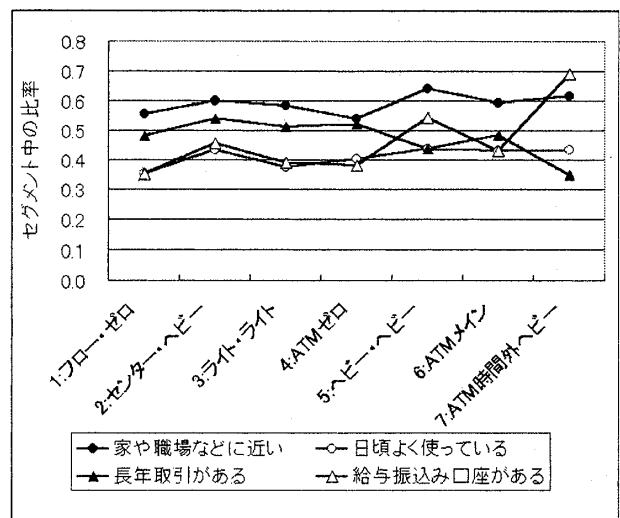


図3 メインバンク選択理由

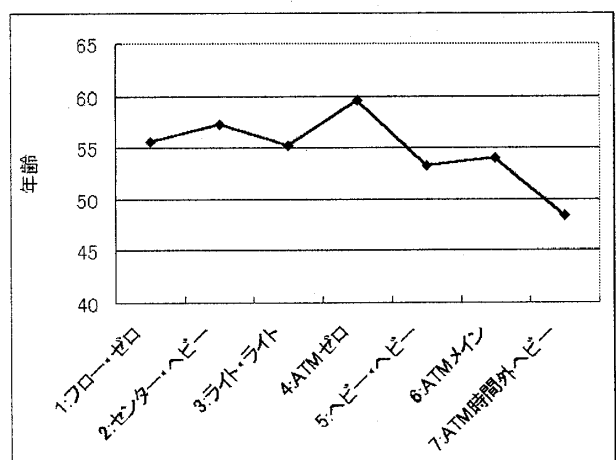


図4 平均年齢

ヘビー」では少ない。

- **平均年齢**：セグメント別の平均年齢を図4に示す。「ATM時間外ヘビー」は平均年齢が低いことが分かる。

4.5 フロー情報からのストック選好の把握

性・年齢や取引ポイント別の管理ではなく、行動の結果であるフローの利用構造（チャンネル別利用頻度をもとにしたセグメンテーション）から、どのようなストック（金融商品）を提供すればよいかを考える。セグメント毎のストック選好が分かれば、チャンネル利用頻度からストック選好を予測することも可能となる。

図5～8はセグメント別のストックへの選好である。これらをまとめると以下ようになる。

- **フロー・ゼロ**：定期預金が中心である。定期満期

時の継続を確保することが大切である。

- **センター・ヘビー**：定期預金が中心である。外貨預金に興味がある。
- **ライト・ライト**：まとまった入金時はとりあえず定期に預ける。預け替えには担当者のすすめが重要となる。MMF・中国ファンドや公社債投信に興味がある。
- **ATMゼロ**：担当者のすすめが貯蓄運用を考え直すときに重要である。株式・社債に興味がある。
- **ヘビー・ヘビー**：金利の変化、まとまった入金（ボーナス）時に貯蓄運用を考え直す。元本保証のない商品全般に興味がある。
- **ATMメイン**：定期満期時に貯蓄運用を考え直す。
- **ATM時間外ヘビー**：保有資産が少ないので積極

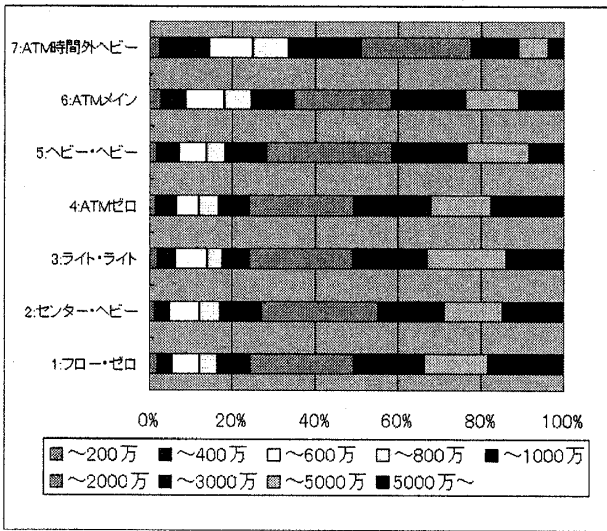


図5 保有資産の状況（回答ベース）

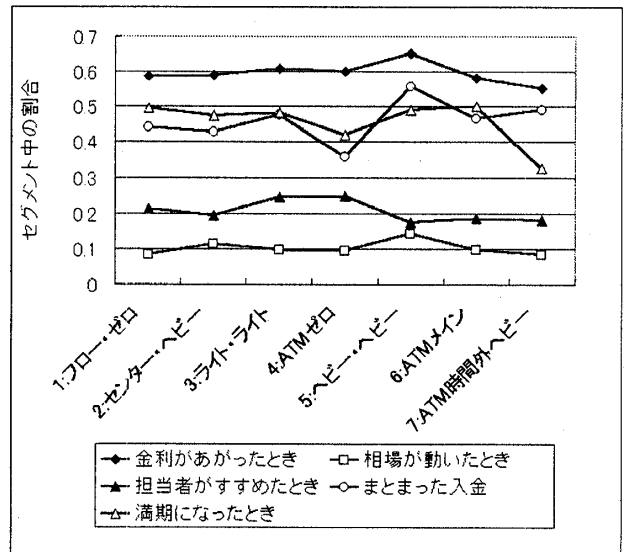


図7 貯蓄運用を考え直すタイミング

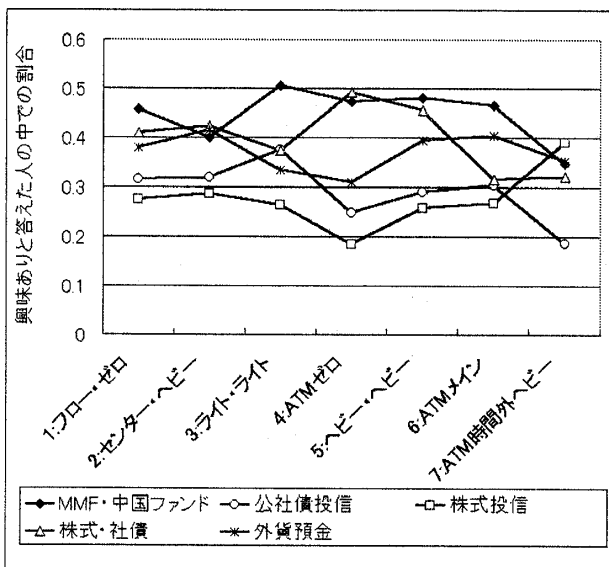


図6 興味のある金融商品（元本保証なし）

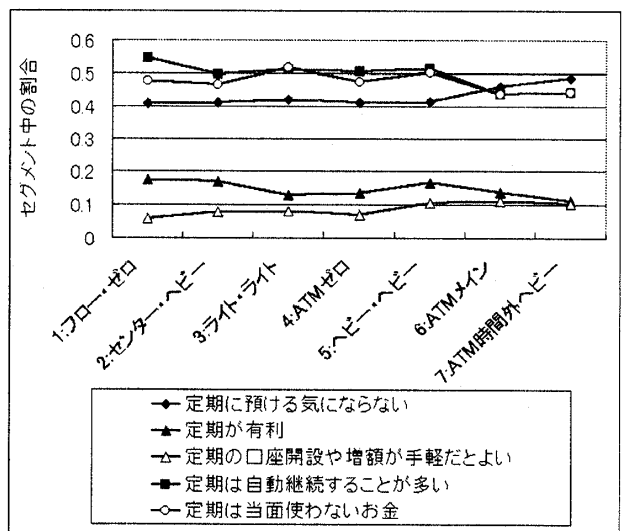


図8 定期予期について

的コミュニケーションの必要はない。ATM時間外利用が多いのでストック商品よりもデビット・カードのような決済機能の商品を提供した方がよいのではないか。

5. おわりに

本研究では、顧客の金融チャネル利用実態から顧客をセグメンテーションする手法について検討を行った。銀行のチャネル利用頻度から顧客をセグメンテーションし、セグメント毎のストックへの選好について見てきた。これらの結果を利用すればチャネル利用頻度からストック選好を予測することも可能となる。

最後に今後の課題について触れてみたい。まずはセグメントへの所属確率をマーケティング変数を用いて説明することである。影響するマーケティング変数が分かればセグメントを移動させるアクションを提案することも可能となる。加えて、顧客が所属するセグメントが時間とともに変化するダイナミック・セグメンテーションを行えば、「行動が変わった顧客は誰か？

行動を変化させるマーケティング変数は何か？」を推定することが可能となる。

DBを利用した個人顧客対応のマーケティングについて今後の発展に期待したい。

貴重なデータの提供をいただいたマーケティング・エンジニアリング研究部会、ならびに多くのご助言をいただいた立教大学の岡太彬訓先生、(株)NTT データシステム科学研究所の中川慶一郎氏をはじめとする研究部会の皆様に感謝いたします。

参考文献

- [1] Dempster, A. P., N. M. Laird and D. R. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM-Algorithm", *Journal of the Royal Statistical Society, Series B* 39, 1-38.
- [2] Wedel, K. and W. A. Kamakura (1998), *Market Segmentation-Conceptual and Methodological Foundations*, Kuwer Academic Publishers.