購買データにおける RFM 指標生成モデルのパラメータ推定に関する研究

情報数理応用研究 5217C027-1 西尾友里
指導教員 後藤正幸

A Study on Parameter Estimation for Generative RFM Model of Purchase Data

NISHIO Yuri

1 研究背景・目的

近年，多くの小売業では，顧客の購買履歴データを解析することで，売上金額増加のための様々な施策に取り組んでいる．従来，購買履歴データから顧客を優良顧客や非優良顧客，生存顧客や離脱顧客などの顧客の特徴を分析するための手法として，RFM 分析 [1] が知られている．この方法は，顧客の購買行動を，R(最終購買からの経過期間)，F(購買頻度)，M(購買金額) の 3 変数によって数量化し，顧客の購買特性を総合的に判断する手法である．RFM 指標に基づき，顧客を企業にとっての重要度でランク付けし，様々なマーケティング施策に結び付けるアプローチは多くの事例で適用されている．

しかし RFM 分析では主に 2 つの課題が存在する．1 つ目の問題点として，顧客の購買特性を分類する際に主観的な要素が強く反映されることが挙げられる．顧客を優良顧客や非優良顧客，また生存顧客や離脱顧客などと特徴付ける際，RFM 指標に着目し，それぞれある閾値に対しての大小関係で識別するが，通常この閾値は，分析者の独自の経験則に基づいて決められてしまう．分析者の判断に依存しないような，より客観的で科学的根拠に基づいた方法が望まれる．2 つ目の課題として，RFM 指標自体からはいつ顧客が離反しそうなのか，いつ来店する可能性が高いのかなど顧客の将来の購買行動に対して具体的な情報を得ることが難しいという点が挙げられる．実際にマーケティング施策を行うことを考えた場合，詳細でかつ将来の購買行動まで把握できるような手法であることが望ましい．

これに対し，Schmittlein ら [2] は，R(最終購買からの経過期間) と F(購買頻度) の指標を用いた顧客行動モデルである Pareto/NBD モデルを提案した．このモデルでは，顧客の購買頻度にポアソン分布，生存時間に指数分布が仮定され，購買頻度と生存時間のパラメータは独立に生起するという仮定を置き顧客の購買行動をモデル化している．さらに阿部 [3] は Pareto/NBD モデルに M の指標を加え，購買金額に対数正規分布を仮定し，購買頻度，生存時間，購買金額に関するパラメータ間の関係を，多変量対数正規分布でモデル化する Hierarchical Bayes モデルを提案した（以下このモデルを阿部の HB モデルと呼ぶ）．これらのモデルは，前述した RFM 分析における 2 つの課題を解消し，分析の際の恣意性の介入を排除し，かつ将来の購買行動を具体的に予測できる方法として知られている．

阿部の HB モデルでは，データから観測されない潜在変数 (顧客の購買行動の観測期間内に顧客は生存中か，または離脱したのかを表す変数と，顧客の生存時間を表す変数) を導入している．そのため MCMC アルゴリズム [4] を導入し，これらの潜在変数と，顧客の購買頻度，生存時間，購買金額に関するパラメータを推定している．このモデルでは，顧客が購買を行う間隔（購買間隔）が分析の対象となっており，観測期間が十分長く取られているような顧客に対しては，潜在変数とパラメータの推定精度は良好で，実データへの当てはまりも良いという特徴がある．観測期間が十分に長いことで観測できる購買の回数が多くなり，観測される RFM データがより顧客の購買嗜好を反映したものとなるためである．しかしながら実際には，家具などの頻繁に購買されないような商品を扱う企業や新規顧客などの存在により，購買間隔に対して観測期間が十分長くとれない顧客が多く，企業が全顧客に対して十分な期間のデータを保持しているケースは稀である．このような購買間隔が長く観測期間が十分取れていない顧客のデータに阿部の HB モデルを適用した場合，特に潜在変数とそれに関わる生存時間に関するパラメータの推定精度が著しく悪化してしまうという問題が生じる．阿部の HB モデルのパラメータ推定アルゴリズムでは，まず潜在変数を推定し，次に潜在変数が既知のもとでの尤度が高くなるようなパラメータを推定するステップが繰り返し行われる．本研究ではまず，従来の推定アルゴリズムについて，観測期間が十分に取れない顧客のほとんどが，観測期間内に顧客が生存中かまたは離脱したのかを表す潜在変数が,「生存」と推定されてしまう特性があることを示す．またこの特性が原因となり，アルゴリズムのステップ数が増加すると，生存時間に関連するパラメータについて大幅に誤った値が推定されてしまう恐れがあることも指摘する．

以上を考慮し，本研究では，阿部の HB モデルのパラメータ推定アルゴリズムを改良し，パラメータに事前分布を仮定し正則化を加えることにより，潜在変数とパラメータの推定精度を改善する手法を提案する．提案手法により，観測期間が十分に取れない顧客の潜在変数やパラメータの推定がよりロバスト行えるようになり，顧客の将来の購買行動をより確からしく把握することが可能となる．また提案した手法を，実際の購買履歴データに適用することでその有用性を示す．

2 阿部の HB モデル

2.1 概要

阿部 [3] は，RFM の 3 指標自体は顧客の購買特性を直接表す指標ではなく，潜在的な購買特性により発生した間接的な指標であると仮定し，RFM 指標の生成モデルを提案している．これは潜在的な購買特性を後に示す 4 つの仮定により表現している．顧客ごとに購買行動をモデル化することで，顧客の将来の購買行動の具体的な把握を可能としている．よって実際のマーケティングにおいて有用な手法であると言える．

2.2 使用する変数

阿部の HB モデルにおいて使用する変数は以下の通りである．いま，ある顧客 $i \in \mathcal{I} = \{1, 2, ..., N\}$ に対し，観測データに記録されている期間のうち，顧客 i の最初に購買した（初期購買）時点を 0，初期購買から顧客 i の最後の購買が発生するまでの期間を t_i，初期購買から観測データが記録されている最後の時点までの期間を T_i としたときの期間 $(0,T_i]$ に発生した購買回数を x_i，また顧客 i の n 回目の購買金額を s_{in} とする．そして，すべての顧客の毎回の購買金額を対数変換した値が従う分布の（全顧客に対する）分散を ω^2 と表記する．

一方，観測データからは観測できない潜在変数として顧客 i の生存時間 τ_i を導入する．生存時間 τ_i は顧客 i の

初期購買が観測された時点 0 から，顧客 i が離脱するまでの期間を表している．阿部の HB モデルでは，生存期間中は購買が発生するが，生存期間を過ぎると顧客は離脱し，それ以降購買は発生しないことが仮定されている．さらに顧客 i が観測期間中に生存しているか離脱しているかを表す変数 z_i も導入しており，z_i は $\tau_i >= T_i$ のとき 1（生存），$\tau_i < T_i$ のとき 0（離脱）を取る指示変数と定義されている．

2.3 阿部の HB モデル

阿部の HB モデルでは顧客の購買特性に対して 4 つの仮定に基づき，各顧客の購買行動を，顧客ごとに特有な顧客特性パラメータで表現する．顧客 i の顧客特性パラメータを $\{\lambda_i, \mu_i, \eta_i\}$ とすると，顧客 i の購買行動は以下の仮定に従うものとする．

仮定 1 顧客 i の購買回数 x_i は，生存期間中，パラメータ λ_i のポアソン分布に従う．

$$P(x_i|\lambda_i) = \begin{cases} \frac{\lambda_i T_i}{x_i!}\mathrm{e}^{-\lambda_i T_i} \ (\tau_i \geq T_i, \text{つまり } z_i = 1 \text{ のとき}) \\ \\ \frac{\lambda_i \tau_i}{x_i!}\mathrm{e}^{-\lambda_i \tau_i} \ (\tau_i < T_i, \text{つまり } z_i = 0 \text{ のとき}) \end{cases} \tag{1}$$

仮定 2 顧客 i の生存時間 τ_i は，パラメータ μ_i の指数分布に従う．

$$f(\tau_i|\mu_i) = \mu_i \mathrm{e}^{-\mu_i \tau_i} \qquad (\tau_i \geq 0) \tag{2}$$

仮定 3 顧客 i の対数変換した購買金額は，パラメータ η_i の正規分布に従う．

$$\log(s_{in}) \sim N(\log(\eta_i), \omega^2) \tag{3}$$

ここで s_{in} は，顧客 i の n 回目の購買金額とする．また ω は過去の情報から既知なものとして扱う．

仮定 4 顧客特性パラメータ $\{\lambda_i, \mu_i, \eta_i\}$ を対数変換した値は多変量正規分布に従う

$$\begin{bmatrix} \log(\lambda_i) \\ \log(\mu_i) \\ \log(\eta_i) \end{bmatrix} \sim MVN\left(\theta_0, \Gamma_0\right) \tag{4}$$

ここで，θ_0，Γ_0 はそれぞれ，多変量正規分布の平均と分散共分散行列を表している．顧客ごとの顧客特性パラメータが推定できれば，個人レベルの購買行動の把握が可能となる．また顧客特性パラメータによって顧客の将来の購買行動を考慮した指標を求めることが可能となる．

2.4 顧客特性パラメータの推定

阿部の HB モデルにおいて使用する観測変数は，顧客 i の RFM 指標つまり，最終購買時点から観測終了時点までの経過時間 $T_i - t_i$(すなわち R)，観測期間中の購買回数 x_i(すなわち F)，観測期間中の平均購買金額 $\overline{s_i}$(すなわち M) である．すべての顧客の毎回の購買金額を対数変換した値が従う正規分布の分散を ω^2 とし，これは過去の情報から既知であるとする．これらの観測変数から顧客ごとに，顧客特性パラメータ $\{\lambda_i, \mu_i, \eta_i\}$ 及び観測されない潜在変数 z_i と τ_i，また顧客全体で仮定される多変量正規分布の平均 θ_0 と分散共分散行列 Γ_0 からなるパラメータ群 $\{\lambda_i, \mu_i, \eta_i, z_i, \tau_i, \theta_0, \Gamma_0\}$ を MCMC 法に基づき推定する．MCMC 法では，各パラメータが，残りのパラメータの値を所与とした条件付確率分布から乱数発生される．モデルの各パラメータは，以下のアルゴリズムを推定値が収束するまで繰り返すことで推定される．

[**MCMC** アルゴリズム]

(Step1) パラメータ群 $\{\lambda_i, \mu_i, \eta_i, z_i, \tau_i, \theta_0, \Gamma_0\}$ の初期値の設定

(Step2) 各顧客 $i = 1, 2, ..., N$ に対して

- **(Step2-a)** $\{z_i|\lambda_i, \mu_i\}$ をサンプリング
- **(Step2-b)** $z_i = 0$ のとき，$\{\tau_i|z_i, \lambda_i, \mu_i\}$ をサンプリング
- **(Step2-c)** $\{\lambda_i, \mu_i|z_i, \tau_i\}$ をサンプリング
- **(Step2-d)** $\{\eta_i|\lambda_i, \mu_i, \theta_0, \Gamma_0\}$ をサンプリング

(Step3) 多変量正規分布のパラメータ $\{\theta_0, \Gamma_0|\lambda_i, \mu_i, \eta_i\}$ を更新

3 提案手法

3.1 提案手法の概要

阿部の HB モデルは，購買間隔に対して観測期間が十分に取れているような顧客に対しては，潜在変数とパラメータの推定精度が高い．しかしながら実際に企業が有するデータでは，家具などの頻繁に購買されないような商品が扱われていることもあり，十分な観測期間が取れないケースも少なくない．また，新たに購買をした新規顧客に対しても十分な観測期間を取ることはできない．そのため購買間隔に対して観測期間が短い顧客に阿部の HB モデルを適用すると，潜在変数とパラメータの推定精度が悪化してしまう．阿部の HB モデルでは，顧客は顧客特有の購買間隔を有していると仮定しており，観測された購買間隔と，直近の購買から観測終了時点までの期間の関係から顧客の生存/離脱を推定しようとしている．そのため特に，直接観測できない生存時間に関わる潜在変数とパラメータの推定精度が著しく悪化してしまう．そこで本研究では，パラメータに事前分布を仮定し正則化を加えることにより既存のパラメータ推定アルゴリズムの精度を改善した手法を提案する．

3.2 阿部の HB モデルにおける推定アルゴリズムの問題点

MCMC アルゴリズムの Step2-a，すなわち顧客の生存離脱を表す潜在変数 z_i の推定ステップに着目する．顧客の生存確率は式 (5) で表され，この確率にしたがって z_i（$z_i = 1$ のとき生存，$z_i = 0$ のとき離脱）がサンプリングされる．

$$P[z_i = 1|\lambda_i, \mu_i, t_i, T_i] = \frac{1}{1 + \frac{\mu_i}{\lambda_i + \mu_i}[\mathrm{e}^{(\lambda_i + \mu_i)(T_i - t_i)} - 1]} \tag{5}$$

次に，λ_i, μ_i を推定するステップ（Step2-c）にも着目する．λ_i, μ_i をサンプリングする事後確率は，尤度関数に対して事前分布が共役でないために直接求められない．そのため独立 MH アルゴリズムによってまず λ_i，次に μ_i をサンプリングする [5]．独立 MH アルゴリズムでは式 (6) の尤度が高くなるような λ_i, μ_i がサンプリングされやすくなる．

$$L(x_i, t_i, T_i|\lambda_i, \mu_i, z_i, \tau_i) \propto \begin{cases} {\lambda_i}^{x_i}\mathrm{e}^{-(\lambda_i + \mu_i)T_i} \\ \\ {\lambda_i}^{x_i}\mu_i\mathrm{e}^{-(\lambda_i + \mu_i)\tau_i} \end{cases} \tag{6}$$

ただし，式 (6) において，上式は $z_i = 1$ のとき，下式は $z_i = 0$ のときを示している．

ここで，阿部の HB モデルにおける問題点を示す．まず，式 (5) の性質として，μ_i に着目すると，値が小さくなるほど生存確率が高くなり，大きくなるほど離脱確率が高くなる．また λ_i に着目すると，値の大きさの単純な比例関係で生存確率は変わらないが，値が小さくなるほど生存確率は μ_i に依存するという特徴がある．一方，式 (6) の性質として，μ_i 及び尤度の関係に着目すると，$z_i = 1$ のとき μ_i が小さくなると尤度が大きくなり，$z_i = 0$ のときは μ_i の大きさの単純な比例関係で尤度の大小が決まらないという点が指摘される．

ここで，購買間隔が長いが観測期間の短い顧客は，観測される購買の回数が少なくなるために λ_i の値が実際よりも小さく推定される傾向にある．そのためこれらの顧客は，上記のようなアルゴリズム中の式の性質より Step2-a の式 (5) で一度 $z_i = 1$(つまり生存) とサンプリングされると，Step2-c では式 (6) より μ_i は，尤度が高くなるよう小さい値がサンプリングされやすくなる．μ_i の値が小さくなると Step2-a の式 (5) で z_i は 1(つまり生存) しかサンプリングされなくなる．そのためステップ数が増加するにつれ μ_i が極小化し，本来は生存でない顧客も生存と判断されやすく，実際の生存時間よりも大幅に生存時間が伸びてしまうという問題が生じる．

3.3　改良推定アルゴリズム

ここでは阿部の HB モデルにおける μ_i の不適切な極小化によって，z_i がうまく推定されないという問題を解決するため，本研究では MCMC アルゴリズム Step2-c の式 (6) の尤度関数の式において，μ_i に正則化を施す方法を提案する．上記で述べた問題は，パラメータを推定する学習データが推定されるパラメータに対して相対的に少ないことで，推定されるパラメータがデータのノイズに大きく影響を受けてしまうことが原因であると考えられる．これは，教師あり学習の分野ではよく知られている問題であり，パラメータが訓練データに適合しすぎることで，未知のデータには適合しない過学習と等価である．

この問題を解決するために，教師あり学習の分野では正則化が広く用いられている．正則化は，教師あり学習における訓練誤差ではなく汎化誤差を最小化するように，パラメータの学習アルゴリズムを改良する方法である．正則化は非常に有力な手法であり，これまで数多くの正則化手法が機械学習のモデルに適用されている．

一方，最適化問題において正則化を行うことは，尤度に対して事前分布を仮定した事後分布を最大化すること等価であることが知られている [6]．言い換えれば，本研究で扱うような教師なし学習の分野においても，パラメータに適切な事前分布を仮定することで，データのノイズの影響を受けずにパラメータを推定することが可能となる．そこで，提案モデルでは，推定がうまくされにくい μ_i に対し事前分布を仮定する．具体的には，Step2-c において，式 (6) の尤度に事前分布で仮定される事前確率を乗じて得られる事後確率から μ_i をサンプリングする．この時，μ_i の事前分布として，μ_i の逆数を引数とし，パラメータ γ を持つ指数分布を仮定する．この事前分布を式 (7) に示す．

$$f\left(\frac{1}{\mu_i}\right) = \gamma \mathrm{e}^{-\gamma \frac{1}{\mu_i}} \tag{7}$$

阿部の HB モデルでは，顧客の生存時間はパラメータ μ_i の指数分布に従うと仮定されていた．すなわち，μ_i の逆数は顧客 i の期待生存時間を示す．従って，提案モデルで導入する事前分布では顧客の期待生存時間が指数分布に従うという仮定をしていると解釈することができ，この仮定は自然なものであると考えられる．さらに，事前分布に指数分布を用いる有効性が過去の研究より確認されていることからも，式 (7) による事前分布の仮定は適切であると考えられる．

したがって，提案手法では，MCMC アルゴリズムの Step2-c において式 (6) の代わりに式 (8) を用いる．

$$L(x_i, t_i, T_i | \lambda_i, \mu_i, z_i, \tau_i) \propto \begin{cases} {\lambda_i}^{x_i} \mathrm{e}^{-((\lambda_i+\mu_i)T_i + \frac{\gamma}{\mu_i})} \\ \\ {\lambda_i}^{x_i} \mu_i \mathrm{e}^{-((\lambda_i+\mu_i)\tau_i + \frac{\gamma}{\mu_i})} \end{cases} \tag{8}$$

ただし，式（8）において，上式は $z_i = 1$ の場合を，下式は $z_i = 0$ の場合を示している．

4　実データ分析

提案手法の有効性を検証するために，株式会社良品計画から提供頂いた購買履歴データに提案手法を適用し，その結果を分析する．

4.1　分析条件

株式会社良品計画から提供された「無印良品」ブランドの購買履歴データに対して提案手法を適用し，その結果として得られる知見について考察する．「無印良品」は，日本で人気の独自ブランドの大手小売り専門店である．一般に，購買履歴データの特徴として，「会員登録した後一回も購買をしない」，または「数回購買しただけで，その後は来店しない顧客が少なくない」などが挙げられる．「無印良品」の購買履歴データも同様な特徴を有していることが確認されている．今回の実験の RFM 指標について，期間の単位は日数，観測期間の中で購買行動が確認された日数を購買頻度，購買金額の単位は 1 万円とした．データの対象期間は 2014 年 3 月 1 日から 2015 年 2 月 28 日までの 1 年間である．対象期間に 1 回以上購買が観測された全ての顧客から，ランダムに選ばれた 1,000 人の顧客を対象顧客とした．

阿部の HB モデルのパラメータの従来の推定手法 (従来手法) と，本研究で提案したパラメータ推定手法 (提案手法) にそれぞれ上記のデータを適用し，その結果を比較する．どちらのモデルにおいても，事前実験の結果より，MCMC アルゴリズムのステップ数は 2,000 回，パラメータの推定に使用しないバーンイン期間は 0 回から 1,000 回と設定し，ω^2 については，顧客の平均購買金額 (M) から得られる推定値を用いた．また阿部の HB モデルでは，顧客特性パラメータの従う多変量対数正規分布のパラメータ θ_0, Γ_0 にさらに事前分布を仮定し，階層ベイズの枠組みを導入している．一方で，本実験では，事前分布を仮定しない代わりに，事前実験より適切な θ_0, Γ_0 の初期値を設定することで対応している．

4.2　分析結果

表 1 は，各モデルから推定された顧客特性パラメータの平均値と分散，またパラメータの値から算出される将来の購買指標値の平均値と分散を示している．

表 1. 各パラメータと各指標の平均値と分散

	従来手法		提案手法	
	平均	分散	平均	分散
λ_i の推定値	8.55	130.40	8.71	122.80
μ_i の推定値	0.05	0.00	0.41	0.02
η_i の推定値	1.72	1.50	1.71	1.46
期待生存時間 (年数)	22.25	16.00	2.68	0.64
1 年後の生存確率	0.95	0.00	0.67	0.01
観測終了時点の生存確率	0.96	0.01	0.81	0.05
1 年後までの期待購買回数	8.35	124.73	7.27	89.45
1 年後までの期待合計購買金額 (万円)	2.28	10.12	1.98	7.40
顧客生涯価値 (万円)	52.65	5623.46	6.87	6.87

表 1 より，2 つのモデル間で，顧客特性パラメータのうち，μ_i の推定値に大きな差異があることがわかる．またパラメータ μ_i の違いによって，将来の購買指標も差異が生じている．特に大きな差異が表れた将来の顧客指標は，「期待生存時間」である．従来手法を適用して得られた結果の期待生存時間の平均値は 22 年であるが，提案手法を適用して得られた結果の平均値は 2.7 年となった．無印良品の顧客の平均生存時間が 22 年であるということは想像し難く，提案モデルで得られた 2.7 年という結果の方が現実的に相応しいと考えられる．提案モデルにおいて μ_i に事前分布を仮定し正則化を施したしたことが，μ_i の過適合を防ぎ現実的な結果を得られたことにつながったと言える．

提案モデルの有効性をより詳細に把握するために，パラメータ μ_i の推定値と将来の購買指標のうち観測終了時点での生存確率に着目する．図 1 は，各モデルの顧客パラメータ μ_i の推定値と人数の度数分布を表したグラフである．この結果からも，従来モデルと提案モデルで得られた μ_i の推定値の分布が異なっており，正則化の効果が確認できる．

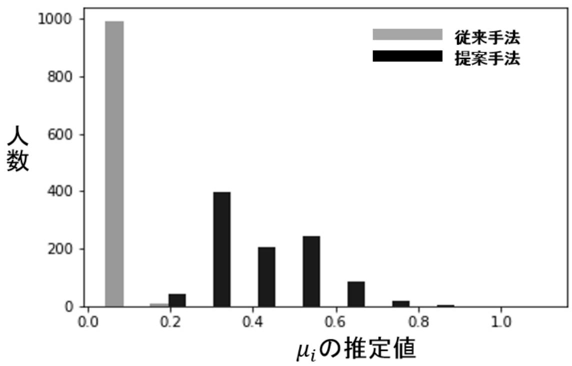

図 1.μ_i の推定値と人数の度数分布

図 2 は，2 つのモデルから得られた，各顧客の観測終了時点の生存確率の値と人数の度数分布を表したグラフである．

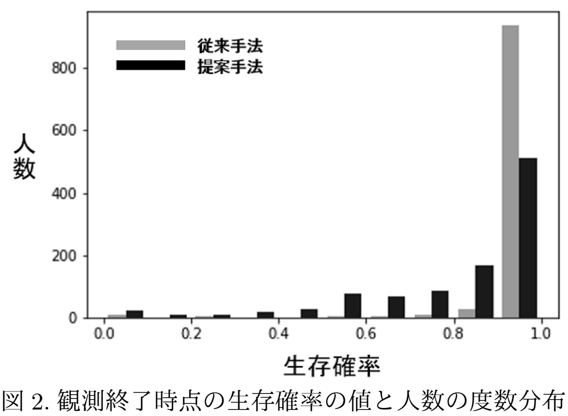

図 2. 観測終了時点の生存確率の値と人数の度数分布

グラフより，従来手法から得られた顧客の生存確率は，ほとんどが 95 パーセント以上であることがわかる．実際のマーケティング施策を行うことを考えると，ほとんどの顧客は離脱することはないので対策を取らなくて良いということになってしまい，ターゲットを絞ることができない．一方，提案モデルでは，離脱しそうな顧客を抽出できており，実際のマーケティング施策を行う顧客ターゲティングに際に有効な結果と言える．

5　考察

購買履歴データには，会員登録後に一度だけ購買しその後二度と購買を行わない顧客のデータが多数含まれる一方，継続的に購買をする顧客も多く存在する．しかしながら，一度きりの購買で離脱してしまう顧客については購買間隔に関する情報が得られていないことから，阿部の HB モデルで，購買間隔が長いのか，あるいは離脱したのかを推定することは難しいと考えられる．本研究ではパラメータに事前分布を仮定し正則化を施したパラメータ推定法により，このような顧客のパラメータ推定に関して阿部の HB モデルを適用した場合の問題点を解決している．今回は全ての顧客の購買データに対してモデルを当てはめたが，一度だけ購買して離脱してしまう顧客を判別した後，継続的に購買する可能性の高い顧客のみに阿部の HB モデルを適用するなど，モデル自体を拡張するアプローチも考えられ，このモデルにはさらなる検討の余地があると考えられる．

また提案手法で仮定した事前分布のパラメータ γ は，正則化パラメータの役割があると考えられる．事前実験により γ の値を変えて阿部の HB モデルに適用し，その時の結果が最も現実と合致していると考えられる値に設定した．今回の実験のようにデータの観測期間が短い場合は過学習が起こりやすいと考えられるので，γ の値を大きくし正則化を強く施すことが望まれるが，パラメータ推定に十分な期間の観測データがある場合には γ の値を小さくし本来の尤度関数に近い関数を用いてパラメータ推定する方が望ましいと考えられる．

6　まとめと今後の課題

本研究では，RFM 指標の生成モデルにおけるパラメータ推定アルゴリズムの改良手法を提案した．提案モデルは，モデルのパラメータに適切な事前分布を仮定する，すなわち正則化を行うことにより，従来よりもロバストにパラメータを推定することを可能とする．特に提案モデルは，十分な期間の購買履歴データを学習データとして用意できないようなケースで有効であり，購買間隔が長い顧客や新規顧客などで十分な期間観測されていない顧客についても，将来の購買行動の予測が可能となる．顧客の将来の購買行動を具体的に把握できるという点で，提案手法は特に実際のマーケティング施策を検討する際に有用な手法であると言える．

今回は正則化のためにパラメータ μ_i の事前分布に指数分布を仮定したが，より適切な事前分布が存在する可能性もあり，この点で検討の余地がある．また，提案手法の有効性の評価について，より解析的に行うことも今後の課題である．

参考文献

[1] Middleton, A.H., “*Strategic database marketing*,”Probus Publishing, Chicago, IL, 1994.

[2] Schmittlein, C.D., Morrison, G.D., and Colombo, R., “Counting your customers: Who are they and what will they do next?” *Management Science*, vol.33, pp.124, 1987.

[3] 阿部誠, “RFM 指標と顧客生涯価値 : 階層ベイズモデルを使った非契約型顧客関係管理における消費者行動の分析,” 日本統計学会誌, vol.41, pp.51-81, 2011.

[4] Gilks, R.W., Richardson,S., and Spiegelhalter, D., “*Markov Chain Monte Carlo in Practice*,” Chapman and Hall/CRC, London, U.K, 1995.

[5] Tierney, L., “Markov Chains for Exploring Posterior Distributions,” *The Annals of Statistics*, vol.22, pp.1701-1762, 1994.

[6] Phillips, J.S., Dudk, M., Schapire, E.R., “A maximum entropy approach to species distribution modeling,” *Proceedings of the Twenty-First International Conference on Machine Learning*, pp.655-662, 2004.