

2013年度 S-PLUS学生研究奨励賞

消費者異質性に着目した 動学的プロビットモデルの構築

筑波大学大学院

システム情報工学研究科 社会システム工学専攻

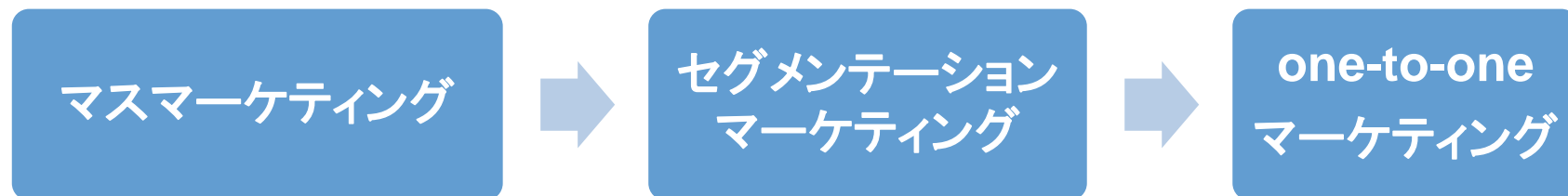
肱岡 浩大

目次

1. 研究背景と目的
2. 研究概要
3. 分析モデル
4. 推定方法
5. モデル評価
6. 使用データ
7. 分析結果
8. まとめと考察
9. 参考文献

研究背景と目的

◆マーケティング形態の推移



◆階層ベイズモデルによる消費者異質性を考慮した分析

➔ 消費者一人一人に対してのアプローチが主流となり、
ベイズモデルによって分析が容易になった

目的

消費者のブランド選択行動に着目し、
ベイズモデルによって複雑な消費者異質性を表現する

研究概要

- ◆ 同研究室の久保田(2010)の研究で使用されているモデルを用いて分析を継続し、研究結果の改善を試みる
- ◆ 久保田(2010)からの改善点(詳細は後述)
 - ◆ モデルの評価方法: DICを用いた評価による最適モデルの決定
 - ◆ 収束判定: Devianceを用いた判定
 - ◆ 状態空間モデルのパラメータ推定方法: FFBSアルゴリズム

ブランド選択モデル

◆ブランド選択行動の表現：確率的効用最大化モデル

…消費者*i*が購買機会*j*にブランド*k*を選択する確率は、
ブランド*k*に対する効用 u_{ijk} が他のすべてのブランドに
対する効用 $u_{ij(-k)}$ より高くなる確率とする

◆効用関数 u_{ijk} の基本的なモデル：

$$u_{ijk} = x_{ijk}\beta + \epsilon_{ijk}, \epsilon_{ijk} \sim N(0, \Lambda)$$

$x_{ijk}\beta$:説明変数によって確定的に決まる要素

ϵ_{ijk} :誤差項

◆ブランド選択結果 y_{ijk} の定義： $u_{ijk} > u_{ij(-k)}$ の時、 $y_{ijk} = 1$

多項プロビットモデル

◆ 効用関数：多項プロビットモデル

- …直接観測されるデータ y_{ijk} が二値(0,1)データであり、正規回帰モデルと潜在変数 u_{ijk} でつながっているモデル

特徴：誤差項に正規分布を仮定.

分散は同じである必要はなく、共分散の表現によって互いに相関を持たせることができる

⇒表現力の豊かなモデルと言われている

同じ二値データを用いた選択モデルの多項ロジットモデルでは、無関係な選択肢からの独立(IIA)という性質がある
(IIA:選択肢の集合が変わっても選択肢間の相対的な選択確率は変わらない)

識別性の問題への対応

- ◆ 多項プロビットモデルには位置と尺度の識別性の問題がある
 - … 全ての選択肢の効用 u_{ijk} に
 - 正の定数を加えた場合でも正の定数を乗じた場合でも
 - ブランド選択確率は同じ結果を示してしまう

- ◆ 位置の識別性への対応 McCulloch and Rossi(1994)

基準ブランド u_{ijK} との効用の差を取り、相対効用 u_{ijk}^* を定義する

$$u_{ijk}^* = u_{ijk} - u_{ijK}$$

- ◆ 尺度の識別性への対応 McCulloch and Rossi(1994)

推定値 β_i を分散の1要素 σ_{11} で割ることで、事後的に識別性を成立させる

$$\beta_i = \beta_i / \sigma_{11}$$

階層ベイズモデル

◆ ベイズモデル:

モデルのパラメータ β 自体に分布を持たせる

$$u_{ijk} = x_{ijk}\beta + \epsilon_{ijk}$$

$$\beta \sim N(\mu_\beta, V_\beta)$$

$$\epsilon_{ijk} \sim N(0, \Lambda)$$

◆ 階層ベイズモデル:

さらに、 β の平均 μ_β ,分散共分散行列 V_β に分布を持たせる

$$\mu_\beta \sim N(\mu_0, \Sigma_0)$$

$$V_\beta \sim IW(\nu_0, \Gamma_0)$$

⇒ 消費者の異質性をモデル化することができる

状態空間モデル

◆ 状態空間モデル:

観測されない状態 Ξ_j によって市場の変動を潜在変数で表現

観測方程式

$$U_{ij} = \beta_i^0 + X_{ij}\beta_i + Z_i\Xi_j + \varepsilon_{ij},$$

→ データとパラメータの関係を示す

システム方程式

$$\Xi_j = F\Xi_{j-1} + Gv_j, v_j \sim N_k(0, Q_\Xi)$$

→ パラメータの時間的推移を記述する

$$U_{ij} = \begin{bmatrix} u_{ij1} \\ \vdots \\ u_{ijk} \end{bmatrix}, \Xi_j = \begin{bmatrix} \xi_{j1} \\ \vdots \\ \xi_{jk} \end{bmatrix}, \varepsilon_{ij} = \begin{bmatrix} \varepsilon_{ij1} \\ \vdots \\ \varepsilon_{ijk} \end{bmatrix}$$

分析モデル

- ◆モデル1 : 基本モデル(状態空間モデル)

$$U_{ij} = \beta_i^0 + X_{ij}\beta_i + Z_i\Xi_j + \varepsilon_{ij}$$

- ◆モデル2 : ブランドロイヤルティを加えたモデル

$$U_{ij} = \beta_i^0 + X_{ij}\beta_i + \log_{ij}\gamma_i + Z_i\Xi_j + \varepsilon_{ij}$$

- ◆モデル3 : γ_i を消費者間で共通のパラメータ γ としたモデル

$$U_{ij} = \beta_i^0 + X_{ij}\beta_i + \log_{ij}\gamma + Z_i\Xi_j + \varepsilon_{ij}$$

- ◆モデル4 : トレードオフパラメータ ϕ_i を加えたモデル

$$U_{ij} = \beta_i^0 + \phi_i X_{ij}\beta_i + (1 - \phi_i)\log_{ij}\gamma + Z_i\Xi_j + \varepsilon_{ij}$$

分析モデル1

◆基本モデル(状態空間モデル)

$$U_{ij} = \beta_i^0 + \underline{X_{ij}\beta_i} + \underline{Z_i\Xi_j} + \varepsilon_{ij}$$

プロモーション
による効用

市場全体の
時間的変化

i:消費者 j:購買機会 k:ブランド(ブランド数K)

U:購買によって得られる効用のK次元ベクトル

β^0 :切片 X:マーケティング変数(説明変数) β :Xの係数パラメータ

Z:ブランドの参入状況 Ξ :状態変数のK次元ベクトル ε :誤差項

分析モデル2

◆モデル1にブランドロイヤルティを加えたモデル

$$U_{ij} = \beta_i^0 + X_{ij}\beta_i + \text{loy}_{ij}\gamma_i + Z_i\Xi_j + \varepsilon_{ij}$$

ブランド・
ロイヤルティ

i:消費者 j:購買機会 k:ブランド(ブランド数K)

U:購買によって得られる効用のK次元ベクトル

β^0 :切片 X:マーケティング変数(説明変数) β :Xの係数パラメータ

loy:ブランド・ロイヤルティ γ :loyの係数パラメータ

Z:ブランドの参入状況 Ξ :状態変数のK次元ベクトル ε :誤差項

ブランド・ロイヤルティの定義

◆ブランド・ロイヤルティloyの定義 Guadagni and Little(1983)

…ブランド選択行動に対するプロモーションの影響を考慮

◆前回購買j-1回目がプロモーション購買かどうかで場合分け

j-1回目の購買がプロモーション購買でない場合:

$$\text{loy}_{ijk} = \pi \text{loy}_{i(j-1)k} + (1 - \pi) y_{i(j-1)k}$$

j-1回目の購買がプロモーション購買である場合(更新しない):

$$\text{loy}_{ijk} = \text{loy}_{i(j-1)k}$$

◆プロモーション購買かどうかの判断

…説明変数(特別価格、特別陳列、チラシ掲載の有無)のうち
2つ以上を伴っている購買をプロモーション購買とする

◆問題点:過去の購買行動によってプロモーション購買か否かを判断

分析モデル3

- ◆モデル2の γ_i を消費者間で共通のパラメータ γ としたモデル

$$U_{ij} = \beta_i^0 + X_{ij}\beta_i + \text{loy}_{ij}\gamma + Z_i\Xi_j + \varepsilon_{ij}$$



i:消費者 j:購買機会 k:ブランド(ブランド数K)

U:購買によって得られる効用のK次元ベクトル

β^0 :切片 X:マーケティング変数(説明変数) β :Xの係数パラメータ

loy:ブランド・ロイヤルティ γ :loyの係数パラメータ

Z:ブランドの参入状況 Ξ :状態変数のK次元ベクトル ε :誤差項

分析モデル4

- ◆モデル3にトレードオフパラメータ ϕ_i を加えたモデル

$$U_{ij} = \beta_i^0 + \phi_i X_{ij} \beta_i + (1 - \phi_i) \text{loy}_{ij} \gamma + Z_i \Xi_j + \varepsilon_{ij}$$

プロモーションによる効用とロイヤルティによる効用のトレードオフパラメータ

i:消費者 j:購買機会 k:ブランド(ブランド数K)

U:購買によって得られる効用のK次元ベクトル

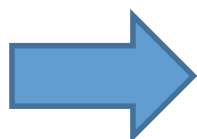
β^0 :切片 X:マーケティング変数(説明変数) β :Xの係数パラメータ

loy:ブランド・ロイヤルティ γ :loyの係数パラメータ ϕ :トレードオフパラメータ

Z:ブランドの参入状況 Ξ :状態変数のK次元ベクトル ε :誤差項

トレードオフパラメータ

- ◆高いブランドロイヤルティを持つ消費者の効用は、価格変化からの影響が小さくなる Krishnamuthi and Papatla(2003)



価格変化以外のマーケティングプロモーションの変数を含めてトレードオフの関係を考慮する

$$U_{ij} = \beta_i^0 + \phi_i X_{ij} \beta_i + (1 - \phi_i) \log_{ij} \gamma + Z_i \Xi_j + \varepsilon_{ij}$$
$$\phi_i = \frac{\exp(\alpha_i)}{\exp(\alpha_i) + 1}$$

プロモーションの効用 $X_{ij} \beta_i$ とロイヤルティの効用 $\log_{ij} \gamma$ に対しパラメータ ϕ_i による重みづけがなされる(実際には α_i を推定)

推定方法(1/2)

- ◆ロイヤルティパラメータ π の推定：グリッドサーチ
 - ・・・久保田(2010)の $\pi = 0.80$ を使用
- ◆状態空間ベクトル $\mathbf{\varepsilon}_j$ パラメータ推定：カルマンフィルタ
 - ・・・1期先予測とフィルタリングのプロセスを繰り返し、
逐次的に $\mathbf{\varepsilon}_j$ の条件付分布を求めるアルゴリズム
(観測方程式とシステム方程式の誤差項が正規分布である性質を利用)

推定方法(2/2)

◆階層ベイズモデルのパラメータ推定：

マルコフ連鎖モンテカルロ法 (MCMC法)

↳ ベイズ推論のサンプリング計算に用いられる手法
初期値の影響が残るバーンイン期間の結果は捨てる

① ギブス・サンプリング

パラメータの完全条件付分布からサンプリングが可能である場合に利用できる

② メトロポリス・ヘイスティングス法

必ずしも完全条件付分布が得られない場合に使用する
(汎用性は高いが、計算効率が悪い)

収束判定方法

- ◆ MCMC法では、パラメータの推定値が定常分布に収束しているか判定する必要がある
- ◆ 判定方法: Gewekeの方法 グラフの目視

サンプルを前半・後半に分け、平均値の差の検定を行う

- ◆ 多くのパラメータを扱う場合、全てのパラメータを収束させるのは難しい
- ⇒ 尤度 (厳密にはDIC計算に用いるDeviance) の収束を見る方法で代用する

モデル評価方法

◆以下の3つの基準によってモデル評価を行う

① DIC (deviance information criteria)

値は小さい方が良い

$$\begin{aligned} \text{DIC} &= \bar{D}(\theta_k) + p_k \\ p_k &= \bar{D}(\theta_k) - D(\bar{\theta}_k) \end{aligned}$$

$\bar{D}(\theta_k)$: MCMC法のiterationの最後に計算される対数尤度 $\times (-2)$ の平均値

$D(\bar{\theta}_k)$: パラメータの事後平均を使用して計算した対数尤度 $\times (-2)$

② インサンプルへのモデル適合率

③ アウトサンプルへのヒット率

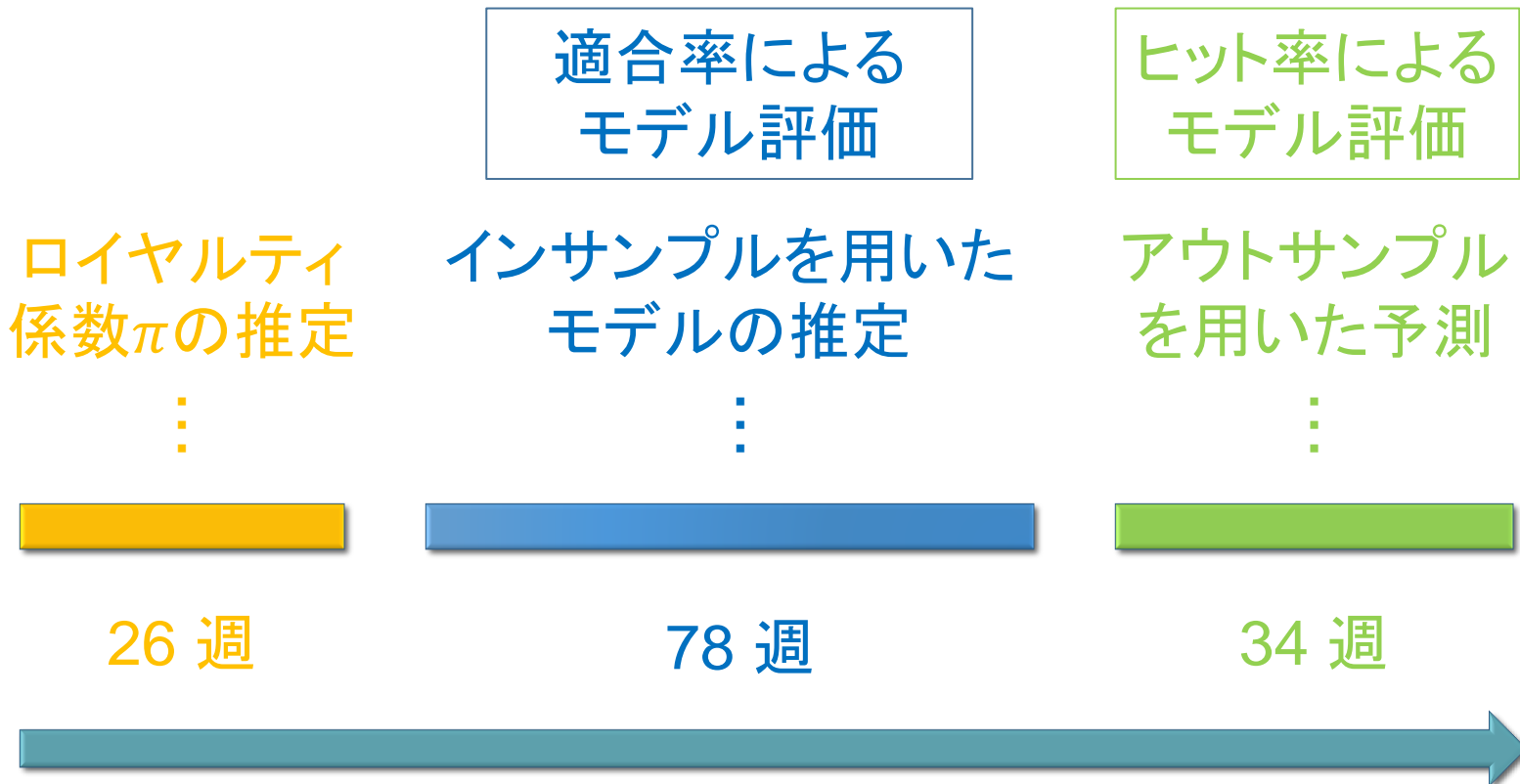
使用データ(1/3)

- ◆ ERIM スキャナー・パネルデータ (商品カテゴリー ketchup)
 - ◆ ブランド数: 4 (全体の67.6%を占める上位4ブランドを選択)
 - ◆ 期間: 1986年第1週～1988年第34週 (計: 138週)
 - ◆ 前処理: 同時に2ブランド以上購買しているデータを除く
 - ◆ 以下の条件を満たすデータに限定
 - ① 最初の26週において、2回以上の購買を行っている世帯
 - ② 次の78週において、5回以上の購買を行っている世帯
 - ③ 最後の34週において、3回以上の購買を行っている世帯

⇒ 条件を満たす311パネルから104パネルをサンプリングし分析対象とする

使用データ(2/3)

- ◆ 期間ごとのサンプルは、それぞれ以下のように使用する



使用データ(3/3)

◆以下の3つの説明変数を用いる

① 割引率 = $1 - \frac{\text{購買価格}}{\text{期間中の最大価格}}$

② 特別陳列の有無 : 0,1のダミー変数

・・・end-aisle display(棚の端の通路に面したディスプレイ)など
購買を促すように特別に陳列された商品

③ チラシ掲載の有無 : 0,1のダミー変数

分析結果

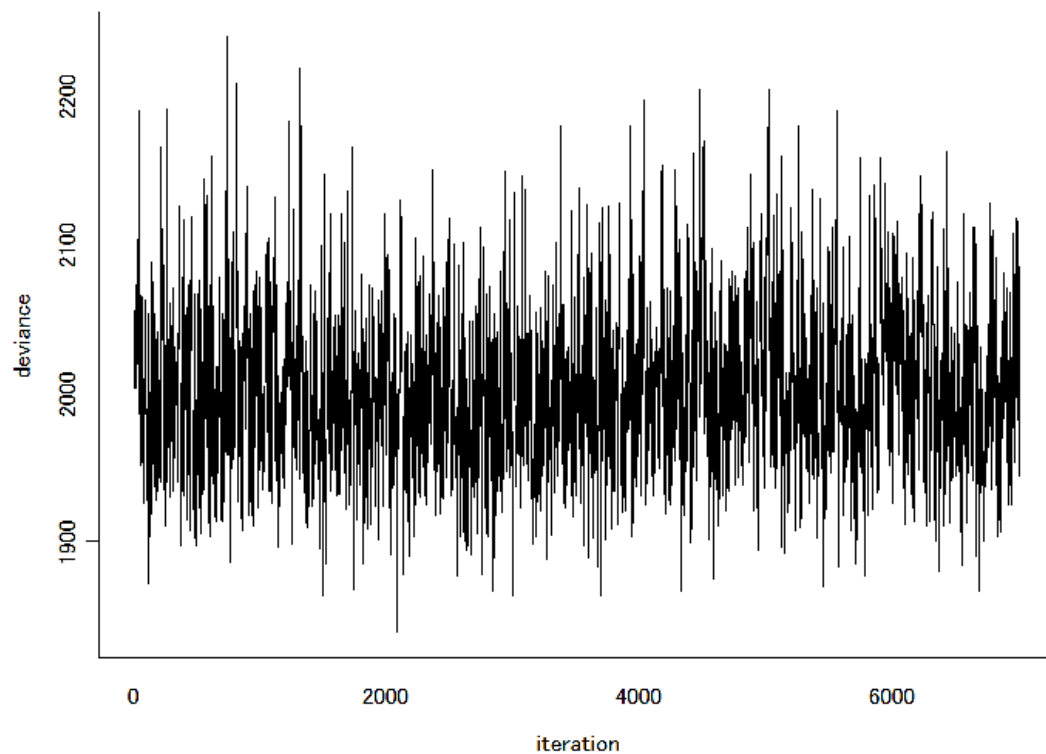
◆各モデルの係数パラメータ値とモデル評価値

	モデル1	モデル2	モデル3	モデル4
割引率	7.935*	8.308*	7.730*	8.987*
特別陳列	0.409*	0.325*	0.548*	0.424*
チラシ	0.375*	0.357*	0.454*	0.399*
ロイヤルティγ	-	1.828*	1.442	1.088
モデル適合率	75.05%	76.15%	77.26%	67.59%
ヒット率	64.01%	67.90%	74.03%	50.46%
DIC	2153	2125	1854	2557

最良値

*:消費者間の平均値

収束判定結果(1/2)



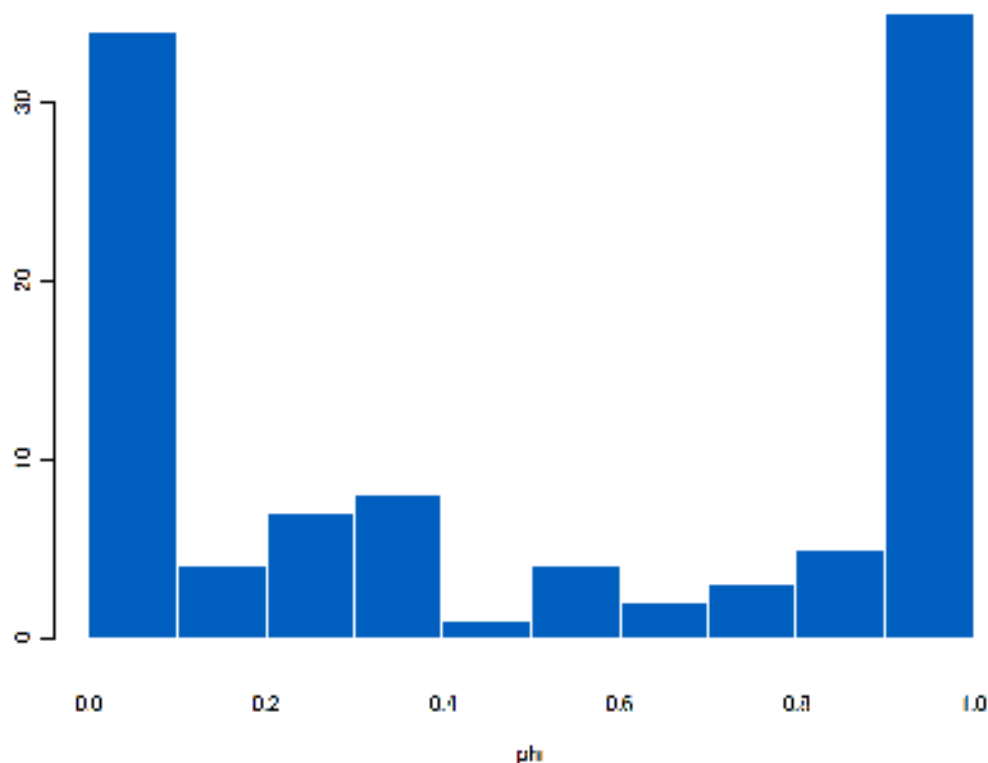
左図のようなグラフの目視と
Gewekeの方法による判定:

モデル1,モデル2,モデル3で
収束を確認

イテレーション回数:14000
バーンイン期間:7000

Devianceの収束確認用グラフ(モデル1)

収束判定結果(2/2)



モデル4における
トレードオフパラメータ ϕ_i

- ・・・消費者ごとに大きな差異が見られた(二分化)

消費者ごとのトレードオフパラメータ ϕ_i のヒストグラム

まとめと考察①

- ◆モデル2ではロイヤルティの係数値が全消費者でほぼ同じ値に推定された
- ◆全ての評価指標でモデル3が最良となっていた
⇒ロイヤルティの係数は消費者ごとに推定する必要がないことが分かった
- ◆モデル1～モデル3でDevianceは収束した
- ◆モデル4は完全に収束しなかったが、トレードオフパラメータの値が消費者ごとに二分化されていることが確認できた
…プロモーションに反応しないロイヤル・カスタマーが存在すると考える

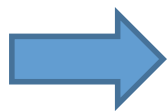
FFBSアルゴリズム

◆FFBS(Forward Filtering Backward Sampling)

Cater and Kohn(1994)

Fruhirth-Schnatter(1994)

- ・・・カルマンフィルタのフィルタリングステップ(forward Filtering)に加え、そこからのサンプリング結果を用いた回顧的なサンプリング(Backward Sampling)を行うアルゴリズム



1期先の情報を用いてサンプリングを行うFFBSによって市場の変動をより正確に表現できるかを検証する

- ・・・モデル1(基本モデル)とモデル3(最良のモデル)を使用

FFBSとの比較

◆各モデルの係数パラメータ値とモデル評価値

	モデル1	モデル1(FFBS)	モデル3	モデル3(FFBS)
割引率	7.935*	8.158*	7.730*	8.056*
特別陳列	0.409*	0.411	0.548*	0.556*
チラシ	0.375*	0.365	0.454*	0.467*
ロイヤルティ γ	-	-	1.442	1.507
モデル適合率	75.05%	69.98%	77.26%	76.15
ヒット率	64.01%	63.67%	74.03%	73.28
DIC	2153	2249	1854	2011

* :消費者間の平均値

まとめと考察②

- ◆モデル1とモデル3共にカルマンフィルタによって推定した方が良い結果となった
 - ⇒本研究のモデルでは、カルマンフィルタによる推定が適していることが分かった
(FFBSの方が実行時間も多く要する)
- ◆FFBSではサンプリングを2度行うため、その分ばらつきが大きくなっている
 - ・・・より多くのイテレーションによって正確な推定を行うことで結果が改善する可能性があると考え

今後の課題

- ◆パラメータの収束判定には、MCMC法のイテレーション回数・バーンイン期間の設定が重要となるため、様々なパターンで試行する
- ◆ロイヤル・カスタマーの存在を考慮することでトレードオフパラメータを設定し直し、モデルを再構築する

参考文献

- [1]久保田郁亮(2010). ブランド・ロイヤルティを考慮した動学的プロビットモデルの構築. 平成22年度筑波大学大学院システム情報工学研究科社会システム工学専攻修士論文.
- [2] Hore, Johannes, Lopes, McCulloch and Polson(2010). Bayesian computation in Finance. *Frontiers of Statistical Decision Making and Bayesian Analysis*,pp.553-578.
- [3]Garrett P.Sonner, Leigh McAlister, Olover J. Rutz(2011). A Daynamic Model of theEffect of Online Communications on Firm Sales. *Marketing Science* 30(4),vol.30,no.4,pp.702-716.
- [4]照井信彦(2010). Rによるベイズ統計分析,朝倉書店.
- [5]古谷知之(2008). ベイズ統計データ分析-R&WinBUGS-,朝倉書店.

Appendix

ギブス・サンプリング

- ◆ 推定するパラメータ m 個の完全条件付事後分布 $f(\beta_t | \beta_{-t}, y)$ からサンプリング可能であるとき、以下の手順に従う

Step1. パラメータの初期値 $(\beta_1^{(0)}, \dots, \beta_m^{(0)})$ を設定し、 $t = 1$ とする

Step2. $l = 1, \dots, m$ に対して以下を発生させる

$$\beta_l^{(t)} \sim f(\beta_l | \beta_{-l}^{(t)}, y)$$

$$\text{ただし、}\beta_{-l}^{(t)} = (\beta_1^{(t)}, \dots, \beta_{l-1}^{(t)}, \beta_{l+1}^{(t)}, \dots, \beta_m^{(t)})$$

Step3. t を $t + 1$ としてStep2に戻る

メトロポリス・ヘイスティングス法

◆ 提案分布 $q(\theta, \vartheta)$ を用いた事後分布 $\pi(\theta|D)$ からのサンプリングは、以下の手順に従う

Step1. $i = 0$ とし、任意の初期値 θ_0 を設定する

Step2. $q(\theta_i, \cdot)$ から候補の値 θ^* 、 $U(0,1)$ から u を発生させる

Step3. $u \leq a(\theta_i, \theta^*)$ であれば $\theta_{i+1} = \theta^*$ 、それ以外は $\theta_{i+1} = \theta_i$

$$a(\theta_i, \theta^*) = \min \left\{ \frac{\pi(\vartheta|D)q(\vartheta, \theta)}{\pi(\theta|D)q(\theta, \vartheta)}, 1 \right\}$$

Step4. $i = i + 1$ とし、Step2に戻る

カルマンフィルタ

- ◆単純な状態空間モデルを仮定する(V_t, W_t が既知のとき)

$$y_t = F'_t \theta_t + v_t, v_t \sim N(0, V_t)$$

$$\theta_t = G_t \theta_{t-1} + w_t, w_t \sim N(0, W_t)$$

- ◆ t 期までの情報を D_t とすると D_t の更新と共に以下のプロセスで逐次的に条件付き分布を求めていく(初期値 m_0, C_0 を与える)

[1]事前分布(予測): $\theta_t | D_{t-1} \sim N(a_t, R_t)$

$$a_t = G_t m_{t-1}, R_t = G_t C_{t-1} G'_t + W_t$$

[2]予測: $y_t | D_{t-1} \sim N(f_t, Q_t)$

$$f_t = F'_t a_t, Q_t = F'_t R_t F_t + V_t$$

[3]フィルタリング: $\theta_t | D_t \sim N(m_t, C_t)$

$$m_t = a_t + A_t (y_t - f_t), C_t = R_t - A_t A'_t Q_t$$

$$A_t = R_t F_t Q_t^{-1}$$

FFBSアルゴリズム

- ◆カルマンフィルタの結果を用いて、以下の手順でサンプリングを行う

Step1.カルマンフィルタのフィルタリング $\theta_n|D_n \sim N(m_n, C_n)$ からサンプリング結果 θ_n を得る

Step2. θ_n を用いて $t = n - 1, n - 2, \dots, 1$ についてサンプリングを行い、 θ_t を得る

$$\theta_t | \theta_{t+1}, D_t \sim N(g_t, H_t)$$

$$g_t = m_t + B_t(\theta_{t+1} - a_{t+1}), H_t = C_t - B_t R_{t+1} B_t$$

$$B_t = C_t G_t R_{t+1}^{-1}$$

Gewekeの方法

◆判定方法は、以下の手順に従う(イテレーション数: R)

Step1. $\beta^{(r)}$ ($r = 1, \dots, R$)を発生させ、前半($r = 1, \dots, s_1$)と後半($r = s_2 + 1, \dots, R$)の平均値を計算する
ただし、通常は $s_1 = 0.1R, s_2 = 0.5R$ とする

$$g_1 = \frac{1}{s_1} \sum_{r=1}^{s_1} g(\beta^{(r)}), g_2 = \frac{1}{s_2} \sum_{r=s_2+1}^R g(\beta^{(r)})$$

Step2. 標準正規分布の Z_α を上側100% α 点、 Z を

$$Z = \frac{g_1 - g_2}{\sqrt{\hat{\sigma}_1^2/s_1 + \hat{\sigma}_2^2/s_2}}$$

として、 $|Z| \leq z_{\alpha/2}$ であれば収束していると判定する