

# マーケティング・リサーチ特論 階層ベイズ多項ロジットモデル

2024年度1学期: 水曜3限  
担当教員: 石垣 司

## 階層ベイズモデル×ブランド選択モデル

### 消費者個人の選択行動の異質性を表現したモデル

- 消費者  $i$  毎に異なる反応係数  $b_i$  を設定し、各消費者のマーケティング活動への反応の異質性をデータから抽出
- 「個」の情報を用いた One to one マーケティングへ
- 本授業: 階層ベイズ多項ロジットモデル



### 【抽出できる情報の例】

佐藤さんはブランドAが200円でB・Cが180円の時のブランドAの選択確率10%  
鈴木さんはブランドAが200円でB・Cが180円の時のブランドAの選択確率90%

## 復習: 多項ロジットモデル #1

### 線形回帰モデルによる効用の確定項の表現

$$V_{itj} = b_{j0} + b_1 x_{itj1} + \dots + b_P x_{itjP}$$

- マーケティング変数ベクトル:  $x_{itj} = [x_{itj1}, \dots, x_{itjP}]^T$
- 反応係数ベクトル:  $b_j = [b_{j0}, b_1, \dots, b_P]^T$
- ブランド価値:  $b_{j0}$

### 多項ロジットモデルによるブランドAの選択確率

$$\begin{aligned} \Pr(Y_{it} = A) &= \frac{\exp(V_{itA})}{\exp(V_{itA}) + \exp(V_{itB}) + \exp(V_{itC})} \\ &= \frac{\exp(V_{itA})}{\sum_{j=1}^J \exp(V_{itj})} \end{aligned}$$

## 階層ベイズ多項ロジットモデル #1

### 観測モデルと事前分布の階層性の関係図

IW: 逆ワイシャート分布  
IG: 逆ガンマ分布

#### 全消費者のパラメータ

$\Phi$  の事前分布

$V$  の事前分布

$$\text{vec}(\Phi) \sim N_{(P \times J)(L+1)}(\text{vec}(\tilde{M}_\Phi), \tilde{\Sigma}_\Phi) \quad V \sim IW(\tilde{\Psi}, \tilde{\nu})$$

#### 各消費者の $i$ の推定対象

$$b_i \text{ の事前分布 } b_i \sim N_{P \times J}(\Phi d_i, V)$$

この破線より上は階層ベイズ線形回帰モデルと同じ構造

$$\text{購買確率 } \Pr(Y_{it} = j) = \frac{\exp(V_{itj})}{\sum_{l=1}^J \exp(V_{itl})}$$

$$V_{itj} = a_{i,j=1} + \dots + a_{i,j=J} + b_{i1} x_{itj1} + \dots + b_{iP} x_{itjP} = \mathbf{x}_{itj}^T \mathbf{b}_i$$

# 階層ベイズ多項ロジットモデル #2

## 確定効用のブランド価値を含んだ表現

- 消費者  $i$  のブランド  $j$  のブランド価値:  $a_{ij}$
- 階層ベイズ多項ロジットモデルの  $b_i$

$$b_i = [a_{ij=1}, \dots, a_{ij=J}, b_{i1}, \dots, b_{iP}]^T$$

$$x_{ijt} = [0, \dots, 1, \dots, 0, b_{i1}, \dots, b_{iP}]^T$$

- 階層ベイズ多項ロジットモデルの確定効用

$$V_{itj} = a_{ij=1} + \dots + a_{ij=J} + b_{i1}x_{itj1} + \dots + b_{iP}x_{itjP} = \mathbf{x}_{itj}^T \mathbf{b}_i$$

- 階層ベイズ多項ロジットモデルの尤度

$$L(\mathbf{y}|\{\mathbf{b}_i\}, \{x_{itj}\}) = \prod_i \prod_t \prod_j \left[ \frac{\exp(\mathbf{x}_{itj}^T \mathbf{b}_i)}{\sum_{l=1}^J \exp(\mathbf{x}_{itl}^T \mathbf{b}_i)} \right]^{y_{itj}}$$

# 階層ベイズ多項ロジットモデルの事後分布の推定 #1

## 多項ロジットモデルの目標分布 $\pi(\mathbf{x})$

- $b_i$  の事前分布:  $b_i \sim N_{P \times J}(\Phi \mathbf{d}_i, V)$
- $b_i$  の事後分布:  $\int p(\{\mathbf{y}_i\}) dy$  は計算困難

$$\pi(\mathbf{b}_i|\Phi, V, \{\mathbf{y}_i\}, \{x_{itj}\}) = \frac{L(\{\mathbf{y}_i\}|\{\mathbf{b}_i\}, \{x_{itj}\}) N_{P \times J}(\mathbf{b}_i|\Phi, V, \{\mathbf{d}_i\})}{\int p(\{\mathbf{y}_i\}) dy}$$

- M-H アルゴリズム(独立連鎖)の採択率  $r$   $\int p(\{\mathbf{y}_i\}) dy$  が消える

$$r = \frac{\pi(\mathbf{b}'_i)q(\mathbf{b}_{i,t}|\mathbf{b}'_i)}{\pi(\mathbf{b}_{i,t})q(\mathbf{b}'_i|\mathbf{b}_{i,t})} = \frac{L(\mathbf{y}|\{\mathbf{b}'_i\}, \{x_{itj}\}) N_{P \times J}(\mathbf{b}'_i|\Phi, V, \{\mathbf{y}_i\}, \{x_{itj}\}, \{\mathbf{d}_i\}) q(\mathbf{b}_{i,t}|\mathbf{b}'_i)}{L(\mathbf{y}|\{\mathbf{b}_{i,t}\}, \{x_{itj}\}) N_{P \times J}(\mathbf{b}_{i,t}|\Phi, V, \{\mathbf{y}_i\}, \{x_{itj}\}, \{\mathbf{d}_i\}) q(\mathbf{b}'_i|\mathbf{b}_{i,t})}$$

#メモ: 多くのベイズモデルの事後分布では  $\int p(\mathbf{y}) dy$  が解析的にも数値的にも計算困難。しかし、M-H法の採択率の計算には事後分布の比のみが必要なため  $\int p(\mathbf{y}) dy$  がキャンセルされて  $r$  の値が計算可能となる。これが、ベイズ推定にMCMCが利用される強力な理由。この理論が普及した1990年代以降、様々な分野でのベイズモデルの応用が爆発的に進んだ

# 階層ベイズ多項ロジットモデルの事後分布の推定 #2

## 階層ベイズ多項ロジットモデルのM-H法

- Pólya-Gamma 分布を用いた提案分布  $q(\mathbf{b}'_i|\mathbf{b}_{i,t})$  (Polson et al. 2013)  
効率的なサンプリングが可能な提案分布  
単純なM-H法と比べて試行錯誤の大幅減
- M-H法とGibbsサンプリングの混合アルゴリズム

For  $t = 0, 1, \dots, T$  まで

$$\begin{cases} \text{For } i = 1, \dots, I \text{ まで} \\ \text{M-Hアルゴリズムで } \mathbf{b}_{i,t+1} \text{ を生成} \\ \text{(以下は線形ベイズ階層帰帰モデルのGibbsサンプラーと同じ)} \\ \hat{\Phi}_{t+1} = (\Delta^T \Delta)^{-1} \Delta^T B \\ \text{vec}(\Phi_{t+1}) \sim N_{(P+1)(L+1)}(\text{vec}\{(\Delta^T \Delta + \tilde{\Sigma}_\Phi)^{-1} (\Delta^T \Delta \hat{\Phi} + \tilde{\Sigma}_\Phi \tilde{M}_\Phi)\}, V \otimes (\Delta^T \Delta + \tilde{\Sigma}_\Phi)^{-1}) \\ V_{t+1} \sim IW(\tilde{\Psi} + (B - \Delta \hat{\Phi})^T (B - \Delta \hat{\Phi}) + (\hat{\Phi} - \tilde{M}_\Phi)^T \{(\Delta^T \Delta)^{-1} + \tilde{\Sigma}_\Phi^{-1}\}^{-1} (\hat{\Phi} - \tilde{M}_\Phi), \tilde{\nu} + I) \end{cases}$$

#メモ: 本授業ではPólya-Gamma分布について詳述しないが、この提案分布のおかげで大規模なモデルでも実用的なパラメータ推定が可能になってきている

# 補足: 多項プロビットモデル

## 多項プロビットモデルによる商品 $j$ の選択確率

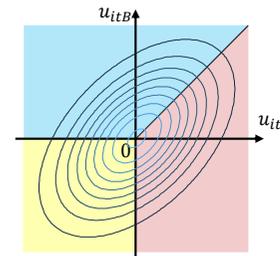
$$\Pr(Y_{it} = j) = \int_{R_{itj}} N_{J-1}(\mathbf{v}_{it}, \Sigma) d\mathbf{u}_{it}$$

相対効用ベクトル:  $\mathbf{u}_{it} = [u_{it1} \ \dots \ u_{it(J-1)}]^T$

相対効用の確定項ベクトル:  $\mathbf{v}_{it} = [v_{it1} \ \dots \ v_{it(J-1)}]^T$

相対効用の分散共分散行列:  $\Sigma$

積分領域  $R_{itj}$  (選択肢  $j$  の相対効用が他の相対効用よりも大きくなる領域)



- 商品 A の選択確率の積分領域 ( $R_{itA}$ )
- 商品 B の選択確率の積分領域 ( $R_{itB}$ )
- 商品 C の選択確率の積分領域 ( $R_{itC}$ )
- $N_{J-1}(\mathbf{v}_{it}, \Sigma)$  の確率密度関数の等高線

#メモ: プロビットモデルの推定には最尤推定、または、ベイズ推定が利用される。しかし、尤度が解析的に計算できないので推定の計算コストが高い

# 補足：階層ベイズ多項プロビットモデル

## 観測モデルと事前分布の階層性の関係図

IW: 逆ウィシャート分布  
IG: 逆ガンマ分布

### 全消費者のパラメータ

Φ の事前分布

V の事前分布

$$\text{vec}(\Phi) \sim N_{(P \times J)(L+1)}(\text{vec}(\tilde{M}_\Phi), \tilde{\Sigma}_\Phi) \quad V \sim IW(\tilde{\Psi}, \tilde{\nu})$$

### 各消費者の i の推定対象

$b_i$  の事前分布  $b_i \sim N_{P \times J}(\Phi d_i, V)$

Σ の事前分布  $\Sigma \sim IW(\tilde{\Sigma}_b, \tilde{\nu}_b)$

この破線より上は階層ベイズ線形回帰モデルと同じ構造

潜在効用  $u_{it} = v_{it} + e_i, e_i \sim N(0, \Sigma)$

データ  $\begin{cases} y_{itj} = 1, (u_{itj} = \max\{u_{it1}, u_{it2}, \dots, u_{it(J-1)}\}) \\ y_{itj} = 0, (\text{otherwise}) \end{cases}$

# 補足：階層ベイズ多項プロビットモデルの推定

## データ拡大法を用いたGibbsサンプリング (Tanner & Wong 1987)

- 相対効用  $u_{it}$  も変数の一つとみなしMCMC法でサンプリング
- $\Phi, V$  は前述のモデルと同じ完全条件付き事後分布
- $\{u_{it}\}, \{b_i\}, \Sigma$  の完全条件付き事後分布を導出可能

## 特徴

- Gibbsサンプリングなので採択効率が良いアルゴリズム
- しかし、逐一、積分  $\int_{R_{itj}} N_{J-1}(v_{it}, \Sigma) du_{it}$  を含む尤度を数値計算する必要があるため計算コストは高い
- 復習：選択肢数の指数乗のオーダーで計算コストが増加

#メモ: 本授業では階層ベイズ多項プロビットモデルの  $\{u_{it}\}, \{b_i\}, \Sigma$  の完全条件付き事後分布の導出は行わないが、興味のある人はチャレンジしてほしい

# Rによる階層ベイズ多項ロジットモデル #1

## “bayesm”パッケージ

- 階層ベイズモデルを扱うパッケージ

## マーガリンの購買履歴データの分析

Allenby, Greg and Peter Rossi (1991), "Quality Perceptions and Asymmetric Switching Between Brands," Marketing Science 10, 185-205.

- 6種類のマーガリンブランド(stick type)  
米国ミズーリ州スプリングフィールドの313世帯で2年間(最低95週)に観測された 1,995 マーガリンの購買履歴(スキャンパネルデータ)
- 説明変数: 価格
- 属性変数: 世帯収入, 家族人数  
(データ内にはブルカー、大卒、定年退職のデータもあるが計算結果が不安定ここでは推定には利用しない)



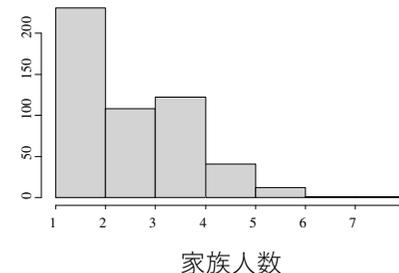
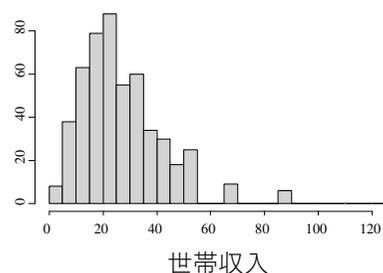
<https://www.fleischmanns.com>  
<https://www.parkay.com/>  
<https://www.gardengrocer.com/product/4587-sheds-spread-country-crock-original-soft-15oz-tub>

# Rによる階層ベイズ多項ロジットモデル #2

## 購買履歴データの記述統計量

	Parkay	Blue Bonnet	Fleischmann's	House Brand	Generic	Imperial
販売個数	1,766	699	243	593	315	74
平均価格(対数)	0.518	0.543	1.015	0.437	0.345	0.781

## 属性データのヒストグラム



# Rによる階層ベイズ多項ロジットモデル #3

## 推定方法

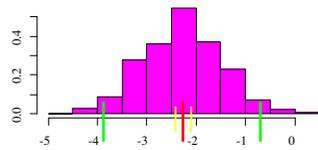
– M-H法(ランダムウォーク連鎖)による推定

ステップサイズ 0.972

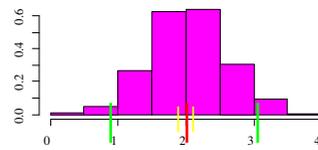
30,000 サンプルを 5 サンプル毎に抽出。Burn-in 期間 1,000 サンプル。合計 5,000 サンプルで事後分布を評価

## 全消費者の $\Phi$ の推定結果

– 家族人数の事後分布(両側 5% 信用区間で有意な結果のみ表示)



House Brand × 家族人数( $\phi_{42}$ )

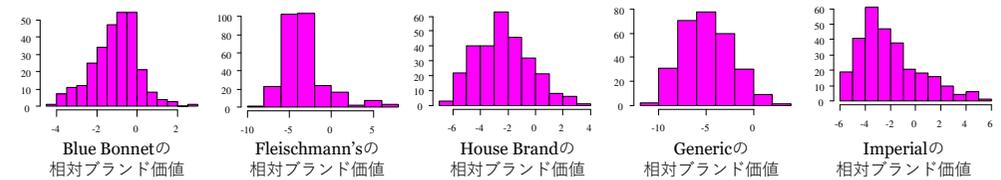


Generic × 家族人数( $\phi_{52}$ )

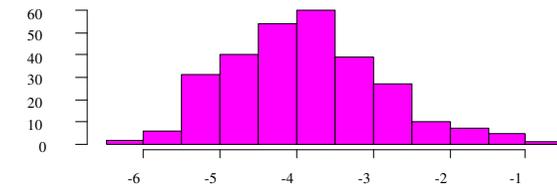
# Rによる階層ベイズ多項ロジットモデル #4

## 消費者の異質性 $\{b_i\}$ の推定結果

– ブランド価値 (各消費者の事後分布の平均値のヒストグラム)



– 価格への反応係数 (各消費者の事後分布の平均値のヒストグラム)



# 応用例の紹介

## 授業内でのみ紹介

# マーケティングと階層ベイズモデリング

照井伸彦, 『ベイズモデリングによるマーケティング分析』, 東京電機大学出版局, 2008

1. 計量経済学では, パネル内の少数データに内在するバイアスを除去して, 経済全体を反映する共通性を推定対象としてきた。異質性はそこでは「バイアス」と呼ばれ, 推定対象外であるので積分して除去される量である。しかし, マーケティングでは, このバイアスこそが貴重な情報であるという逆転の発想があり, ここに大きな特徴がある。
2. 経済学では意思決定者や分析者の主観的判断を極力排除しようとするのに対して, マーケティングでは経験やビジネスセンス, 経営手腕としての主観的な判断を意思決定に積極的に取り入れる土壌がある。そして階層ベイズモデルは, 異質性と共通性の中で情報量の振り分けを自動的に行うというメリットがある。

# 本授業についてのまとめ

---

## **マーケティングの意義は社会・市場での価値の創出**

- 単なる市場調査・製品管理よりもその範囲と意義は広い

## **本授業ではマーケティング・リサーチで利用される定石的なデータ分析手法を紹介**

- 定石とは先人たちの創意工夫のエッセンスを抽出した内容
- 実利用のためには、定石に基づいて自らの創意工夫が必要

## **マーケティングにおけるデータ活用は現代的な課題**

- その方法は社会環境や科学技術の普及で変化する  
必要な知識は時代に合わせてアップデートが必要  
学際的(経営, 経済, 数学, コンピュータなど)な知識が必要