離散選択モデルの導入

安田　公大

一橋大学経済学研究科

作成日: 2023 年 10 月 15 日
最終更新日: 2024 年 5 月 1 日

目次

自己紹介

- 名前: 安田 公大
- 専攻: 実証産業組織論
  - 関心分野: シェアリング・エコノミー、プラットフォーム
- 副専攻: ネットワーク分析
- 学歴:
  - 一橋大学経済学研究科修士課程（2024 - 現在）
  - 一橋大学経済学部（2020 - 現在）
- 職歴:
  - RA（経済学研究科）, 2023 - 現在
  - RA（イノベーション研究センター）, 2023 - 現在

目次

実証産業組織論における分析手法

- 巷で用いられるデータ分析手法 $\neq$ 実証産業組織論での分析手法
- 誘導形アプローチ
  - 因果推論的アプローチとも
  - 回帰や DID、RDD といった因果推論の手法
  - ビジネスでの「データ分析」
- 構造推定アプローチ
  - 経済モデルを構築し、それに対して推定を行う手法
  - 高度な経済学の理解が必要
  - 誘導形にはできない分析が可能
- 本連載では構造推定の諸手法を学んでいく
  - 本日は、全ての根幹となる（静的）離散選択モデルを扱う

構造推定アプローチの長所・短所

- 長所
  - 経済主体の意思決定が重要な状況を分析できる
    - 経済主体の相互依存的関係
    - 経済主体の動学的状況
  - 反実仮想シミュレーションを行うことができる
    - 実際には観測されなかった事象の効果の計測
  - 経済学に関するインプリケーションを得やすい
    - 利潤や厚生（余剰）

構造推定アプローチの長所・短所

- 長所
  - 経済主体の意思決定が重要な状況を分析できる
    - 経済主体の相互依存的関係
    - 経済主体の動学的状況
  - 反実仮想シミュレーションを行うことができる
    - 実際には観測されなかった事象の効果の計測
  - 経済学に関するインプリケーションを得やすい
    - 利潤や厚生（余剰）
- 短所
  - 経済モデルには強い仮定が必要
  - 分析が煩雑になりやすい
    - とにかく時間と高度な専門知識が必要
    - 第三者への平易な説明が難しい
    - ビジネスで使われない（使えない）要因

目次

目次

離散選択とは

- 複数ある選択肢の中から、どれかを選ぶこと
  - お～いお茶、綾鷹、生茶から、お～いお茶を購入する
  - 「今日買う」、「明日以降買う」から、「今日買う」を選択する
  - 「買わない」という選択肢もあり、アウトサイドグッズ（外部財）という
- 選択肢にはそれぞれ特徴があり、経済主体は好みを持つ
  - 味
  - 内容量
  - 価格
- 経済主体は各々の効用を最大化するように選択を行う（顕示選好）

セットアップ

表記は上武 et al. (2021a) に準拠

モデル

- 消費者 $i = 1, \cdots, N$
- 選択肢 $j = 0, \cdots, J \quad (j = 0$は外部財$)$
- 間接効用関数 U_{ij}:

$$U_{ij} = \begin{cases} \beta^0 + \sum_{k=1}^{K} \beta^k x_j^k - \alpha p_j + \epsilon_{ij} & j = 1, \cdots, J \\ \epsilon_{i0} & j = 0. \end{cases} \tag{1}$$

  - 製品 j の価格 p_j
  - 製品 j の製品属性 $x_j^k \quad (k = 1, \cdots, K)$
  - 選好パラメータ $\alpha, \beta^0, \beta^1, \cdots, \beta^k$
  - 個人 i・製品 j 特有のランダムショック ϵ_{ij}

- 消費者は自らの効用 U_{ij} が最大化されるように選択を行う

$$d_i = \underset{j \in \{1, \cdots, J\}}{\arg\max}\; U_{ij}$$

離散選択と選択確率

- 消費者にとって、選択は決定論的になる
  - ある消費者の効用関数が次のようになっているとする

$$U_{ij} = 2 + 2 \times (\text{苦み}) - 3 \times (\text{甘み}) + 3 \times (\text{内容量}) - 2 \times (\text{価格})$$

  - このとき、各メーカーのお茶の効用を計算できる

	苦み	甘み	内容量 (100ml)	価格 (100 円)	効用
お〜いお茶　濃い茶	2	0	5	1.1	18.8
綾鷹　茶葉のあまみ	1	3	5	1.2	7.6
生茶	1	1	5.25	1	14.75

  - 消費者はお〜いお茶　濃い茶を購入する
- 分析者は消費者の選択を確実には予想できない
  - 選好ショック ϵ_{ij} が存在するため
  - 消費者の選択を確率的に求める

$$Pr(d_i = j) = Pr(\{\epsilon_{ij}\}_{j=0}^{J} \ s.t. \ U_{ij} > U_{il} \ \forall l \neq j)$$

  - これを（消費者 i による財 j の）購入確率あるいは選択確率という

目次

多項ロジットモデル（MNL）の選択確率

- 選択確率は $\{\epsilon_{ij}\}_{j=0}^{J}$ の分布を仮定する必要がある

$$Pr(d_i = j) = Pr(\underbrace{\{\epsilon_{ij}\}_{j=0}^{J}}_{\text{この部分}} \ \ s.t. \ U_{ij} > U_{il} \ \ \forall l \neq j)$$

- 多項ロジットモデルでは $\{\epsilon_{ij}\}_{j=0}^{J}$ が $i.i.d.$ に第 1 種極値分布に従うと仮定

$$Pr(\epsilon_{ij} \leq x) = F(x) = e^{-e^{-x}}$$

- このとき、次の選択確率が導かれる

$$Pr(d_i = j) = \frac{\exp(\beta^0 + \sum_{k=1}^{K} \beta^k x_j^k - \alpha p_j)}{1 + \sum_{j=1}^{J} \exp(\beta^0 + \sum_{k=1}^{K} \beta^k x_j^k - \alpha p_j)}$$

  - 分母の 1 は外部財（付録を参照）
  - $\beta^0 + \sum_{k=1}^{K} \beta^k x_j^k - \alpha p_j$ を確定項と呼び、V_{ij} と表記する
  - 確定項が大きくなればなるほど選択確率は上昇する

モデルの推定：最尤推定法

- 必要なデータ $\cdots$ 外部から観察可能
  - 財の特性や価格 $\{p_j, x_j^1, \cdots, x_j^K\}_{j=1}^J$
  - 消費者の選択 $\{d_i\}_{i=1}^N$
- 最尤推定法 $\cdots$ 尤度関数を最大化するパラメータを推定する

$$L(\theta) = \prod_{i=1}^{N}\prod_{j=0}^{J}(Pr(d_i = j))^{\mathbb{1}(d_i=j)}$$

  - 一般には対数尤度を用いる
- ロジットモデルの期待効用は次のように計算できる

$$E_{\{\epsilon_{ij}\}_j}\left[\max_{j\in\{0,1,\cdots,J\}} U_{ij}\right] = log\left(\sum_{j=0}^{J}\exp(V_{ij})\right) + C$$

  - V_{ij} は確定項、$C = 0.57721$ はオイラー定数
  - 導出は McFadden (1977) を参照

選択確率の実践

- 先ほどのお茶の例を用いて、選択確率や期待確率を計算する．「お～いお茶　濃茶」を購入する確率は

$$
\begin{aligned}
Pr(d_i = 1) &= \frac{\exp(\beta^0 + \sum_{k=1}^{K} \beta^k x_1^k - \alpha p_1)}{1 + \sum_{j=1}^{J} \exp(\beta^0 + \sum_{k=1}^{K} \beta^k x_j^k - \alpha p_j)} \\
&= \frac{\exp(18.8)}{1 + (\exp(18.8) + \exp(7.6) + \exp(14.75))} \\
&= 0.9829. \quad (= 98.29\%)
\end{aligned}
$$

- 他の選択肢 $j = 0, 2, 3$ についても同様に計算する．期待効用は

$$
\begin{aligned}
E_{\{\epsilon_{ij}\}_j}\left[\max_{j \in \{0,1,\cdots,J\}} U_{ij}\right] &= log\left(\sum_{j=0}^{J} \exp(V_{ij})\right) + C \\
&= 18.82 + C.
\end{aligned}
$$

目次

MNL モデルの問題点：IIA 特性

- IIA 特性（Independence from Irrelevant Alternatives）
- 財 l, m の選択確率の比が財 l, m 以外の財によらず一定である性質

$$
\begin{aligned}
\frac{Pr(d_i = l)}{Pr(d_i = m)} &= \frac{\exp(V_{il})}{1 + \sum_{j=1}^{J} \exp(V_{ij})} \Bigg/ \frac{\exp(V_{im})}{1 + \sum_{j=1}^{J} \exp(V_{ij})} \\
&= \frac{\exp(V_{il})}{\exp(V_{im})}
\end{aligned}
$$

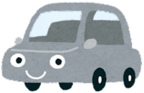
IIA 特性：赤バス - 青バス問題

- 初めに、消費者が自家用車と青いバスから選択する状況を考える

- 選択確率: 0.5

- 選択確率: 0.5

- 次に、赤バスが新たに選択肢に加わる



- IIA: 0.33
- 妥当: 0.5

- IIA: 0.33
- 妥当: 0.25

- IIA: 0.33
- 妥当: 0.25

- このように、IIA 特性が導く選択確率は妥当性に欠けることがある

IIA 特性が招く問題の原因

IIA 特性の性質

- 選択肢の集合によらず、選択確率の比が一定
- 選択肢同士の相関（類似度）を捨象している
- 仮に選択肢同士に相関があると誤差項の分布に矛盾する

交差価格弾力性

- 選択肢間の相関を捨象すると、交差価格弾力性が一定になる

IIA 特性の解決

- IIA 特性を緩和（解決）できるモデルとしては以下が知られる
  - 多項プロビットモデル（MNP）
  - 一般化極値モデル（GEV）
    - 入れ子ロジットモデル（NL）
    - クロス入れ子ロジットモデル（CNL）
    - 不均一分散極値モデル（HEV）
  - 混合ロジットモデル（ML, MXL）
  - ランダム係数ロジットモデル（RCL）
- 他にも、ベイズ推定を用いたアプローチもある

多項プロビットモデル（MNP）

表記は Train (2009) に準拠

- 多項プロビットモデルでは $\{\epsilon_{ij}\}_{j=1}^{J}$ が平均 0、分散 $\mathbf{\Sigma}$ の多変量正規分布に従うと仮定

$$\phi(\tilde{\epsilon}_{ij}) = \frac{1}{(\sqrt{2\pi})^{J-1}|\tilde{\mathbf{\Sigma}}_j|^{1/2}} e^{-\frac{1}{2}\tilde{\epsilon}'_{ij}\mathbf{\Sigma}_j^{-1}\tilde{\epsilon}_{ij}}.$$

where

$$\tilde{\epsilon}_{ij} = \{(\epsilon_{i1}-\epsilon_{ij}),\cdots,(\epsilon_{i(j-1)}-\epsilon_{ij}),(\epsilon_{i(j+1)}-\epsilon_{ij}),\cdots,(\epsilon_{iJ}-\epsilon_{ij})\}$$
$$\mathbf{\Sigma}_j^{-1} = Cov(\tilde{\epsilon}_{ij}).$$

- このとき、次の選択確率が導かれる

$$Pr(d_i = j) = \int_{\epsilon_{i1}=-\infty}^{\epsilon_{ij}+V_{ij}-\epsilon_{i1}} \cdots \int_{\epsilon_{ij}=-\infty}^{\infty} \cdots \int_{\epsilon_{iJ}=-\infty}^{\epsilon_{ij}+V_{ij}-\epsilon_{iJ}} \phi(\tilde{\epsilon}_{ij})d\tilde{\epsilon}_{ij}.$$

  - $(J-1)$ 重積分になり、パラメータ推定が非常に困難になる
  - 計算能力や推定アルゴリズムが向上する近年まで研究の進歩がなかった

入れ子ロジットモデル（NL）：仮定

表記は Cardell (1997), Wakamori (2022) に準拠

- 以後、効用関数 U_{ij} (1) を行列形式で表現する

$$U_{ij} = \begin{cases} \underbrace{\mathbf{X}_j'\boldsymbol{\beta} - \alpha p_j}_{V_{ij}} + \epsilon_{ij} & j = 1, \cdots, J \\ \epsilon_{i0} & j = 0. \end{cases}$$

NL モデル（Two-level Nested Logit モデル）

- いま、財 $1, \cdots, J$ が $G+1$ 個のグループに分けられると想定する
  - $\mathcal{G}_g$: グループ g に含まれる財の集合（$g = 0, 1, \cdots, G$）
  - グループ 0: 外部財の単集合．すなわち、$\#\mathcal{G}_0 = 1$
  - $\bigcup_{g=0}^{G} \mathcal{G}_g = \{1, \cdots, J\}$
- グループ内での相関を認めるが、グループ間での相関は認めない

NL モデル: セットアップ

NL モデル

$$U_{ij} = \begin{cases} V_{ij} + \xi_j + \xi_{ig} + (1-\sigma)\epsilon_{ij} & j = 1, \cdots, J \\ \xi_0 + \xi_{i0} + \epsilon_{i0} & j = 0. \end{cases}$$

where

- ξ_j: 財 j 特有の需要に影響する要因
  - **観測できない**財 j の製品属性
  - 財 j のブランド力
- ξ_{ig}: グループ g 特有の需要に影響する要因
- σ: グループ内の財同士の相関の強さを表すパラメータ
  - $\sigma \to 1$: グループ内の財が相関
  - $\sigma \to 0$: グループ内の財が独立

- このとき、財 $j \in \mathcal{G}_g$ の選択確率から財 j のグループ g におけるマーケットシェアを得る

NL モデル: マーケットシェアの導出

- 財 j の選択確率 $s_j(V_j,\sigma)$ は以下の通り

$$
\begin{aligned}
s_j(V_j,\sigma) &= s_{j/g}(V_j,\sigma)s_g(V_j,\sigma) \\
&= \frac{\exp(V_j/(1-\sigma))}{\sum_{j\in\mathcal{G}_g}\exp(V_j/(1-\sigma))} \times \frac{(\sum_{j\in\mathcal{G}_g}\exp(V_j/(1-\sigma)))^{1-\sigma}}{\sum_{k=0}^{G}(\sum_{j\in\mathcal{G}_k}\exp(V_j/(1-\sigma)))^{1-\sigma}} \\
&= \frac{\exp(V_j/(1-\sigma))}{(\sum_{j\in\mathcal{G}_g}\exp(V_j/(1-\sigma)))^{\sigma}[\sum_{k=0}^{G}(\sum_{j\in\mathcal{G}_k}\exp(V_j/(1-\sigma)))^{1-\sigma}]}.
\end{aligned} \tag{2}
$$

- グループ内での IIA 特性を確認する．グループ g に属する財 j,m の選択確率の比は (2) の分母が打ち消し合うから、

$$
\frac{s_j(V_j,\sigma)}{s_m(V_m,\sigma)} = \frac{\exp(V_j/(1-\sigma))}{\exp(V_m/(1-\sigma))}.
$$

NL モデル: まとめ

NL モデルの特徴

- NL モデルは GEV モデルの一種
  - 限定的に相関を認めてはいるが、誤差項は依然として第 1 種極値分布に従う
  - 長所と短所
    - 長所: closed-form なのでシミュレーションに頼らず選択確率が計算できる
    - 短所: GEV モデルのバリエーションが非常に少ない / 近年ではシミュレーション技術が発達してきた
- NL モデルの拡張としてはクロス入れ子モデル（CNL）やツリー型極値（TEV）モデルが知られる

NL モデルの問題点

- 推定の前に階層構造を明確化する必要がある（難しい）
- 設計した階層構造に誤りがあるとパラメータが正しく推定できなくなる

ランダム係数ロジットモデル（RCL）

表記は Wakamori (2022), 上武 et al. (2021b) に準拠

- 仮定: 製品属性 $\mathbf{X}_j$ や価格 p_j から受ける効用は消費者間で同質
  - しかし、消費者によっては製品属性や価格に対する感度が異なる
- パラメータを個人ごとに設定 $\cdots$ ランダム係数ロジットモデル

$$U_{ij} = \begin{cases} \mathbf{X}_j' \boldsymbol{\beta}_i - \alpha_i p_j + \xi_j + \epsilon_{ij} & j = 1, \cdots, J \\ \boldsymbol{\epsilon}_{i0} & j = 0. \end{cases}$$

  - RCL モデルは、NL モデルの特殊なケースとして捉えることができる
  - また、RCL モデルは MXL モデルの特殊なケースとして捉えることもできる
    - 詳細は付録を参照

RCL モデル: 効用関数の変形

- 効用関数を変形する

$$
\begin{aligned}
U_{ij} &= \mathbf{X}_j'\boldsymbol{\beta}_i + \alpha \log(y_i - p_j) + \xi_j + \epsilon_{ij} \\
&= \sum_{k=1}^{K} x_{j,k}\beta_{i,k} + \alpha \log(y_i - p_j) + \xi_j + \epsilon_{ij} \\
&= \sum_{k=1}^{K} x_{j,k}(\bar{\beta}_k + \sigma_k \nu_{i,k}) + \alpha \log(y_i - p_j) + \xi_j + \epsilon_{ij} \\
&= \underbrace{\sum_{k=1}^{K} x_{j,k}\bar{\beta}_k + \xi_j}_{\delta_j} + \underbrace{\sum_{k=1}^{K} x_{j,k}\sigma_k \nu_{i,k} + \alpha \log(y_i - p_j)}_{\mu_{ij}} + \epsilon_{ij} \\
&= \delta_j + \mu_{ij} + \epsilon_{ij}
\end{aligned}
$$

where

  - $\nu_{i,k}$: 観察者にランダムな効用のショック. $\nu_{i,k} \sim N(0,1)$
- U_{ij} は消費者間で同質な項 δ_j と消費者に得意な項 μ_{ij} に分解できる
  - δ_j を消費者の平均効用という
  - μ_{ij} は消費者 i の平均効用からの乖離を表す

マーケットシェアの導出

- ϵ_{ij} は $i.i.d.$ に第 1 種極値分布に従うと仮定しているので、財 j の選択確率が求まる

$$Pr(d_i = j \mid \{\nu_{ik}\}_{k=1,\cdots,K}) = \frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_{l=1}^{J} \exp(\delta_l + \mu_{il})}.$$

- 財 j を選択する消費者の集合は以下のようになる:

$$\begin{aligned} A_j &= \left\{ \left(\{\epsilon_{ih}\}_{h=1}^{J}, \{\nu_{ik}\}_{k=1}^{K} \mid u_{ij} \geq u_{il}, \forall l \neq j \right) \right\} \\ &= \left\{ \left(\{\epsilon_{ih}\}_{h=1}^{J}, \{\nu_{ik}\}_{k=1}^{K} \mid \epsilon_{il} \leq (\delta_j - \delta_l) + (\mu_{ij} - \mu_{il}) + \epsilon_{ij}, \forall l \neq j \right) \right\} \end{aligned}$$

- したがって、財 j の市場全体におけるマーケットシェアが求まる

$$s_j = \underbrace{\int \cdots \int}_{\nu} \Pi_{l \neq j} \left[\frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_{l=1}^{J} \exp(\delta_l + \mu_{il})} \right] dF_{\nu_1}(\nu_{i1}) \cdots dF_{\nu_K}(\nu_{iK}).$$

- しかし、これの Berry Inversion を解析的に導出することはできない

Cardell, N. Scott (1997). “Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity”. In: *Econometric Theory* 13.2, pp. 185–213. URL: http://www.jstor.org/stable/3532724 (visited on 10/14/2023).

McFadden, Daniel (1977). “Modelling the choice of residential location”. In.

Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. 2nd. Cambridge Books 9780521766555. Cambridge University Press, Dec. 2009.

Wakamori, Naoki (2022). *Advanced Industrial Economics I Class #03, #05, #07*.

上武, 康亮, 祐太 遠山, 直樹 若森, and 安虎 渡辺 (2021a). “実証ビジネス・エコノミクスとは”. In: vol. 1. 実証ビジネス・エコノミクス. 日本評論社.

— (2021b). “プライシングの真髄は代替性にあり”. In: vol. 3. 実証ビジネス・エコノミクス. 日本評論社.

目次

MNL モデルの選択確率の導出

- 多項ロジットモデルの選択確率を導出する[1]
- はじめに、第 1 種極値分布の密度関数、累積分布関数は以下の通り

$$f(\epsilon_{ij}) = e^{-\epsilon_{ij}} e^{-e^{-\epsilon_{ij}}}, \quad F(\epsilon_{ij}) = e^{-e^{-\epsilon_{ij}}}$$

- ここで、消費者 i による財 j の選択確率は以下のようになる

$$\begin{aligned} Pr(d_i = j) &= \int \left(\Pi_{k \neq j} e^{-e^{-(\epsilon_{ij}+V_{ij}-V_{ik})}} \right) e^{-\epsilon_{ij}} e^{-e^{-\epsilon_{ij}}} d\epsilon_{ij} \\ &= \frac{e^{V_{ij}}}{\sum_{k=0}^{J} e^{V_{ik}}} \\ &= \frac{e^{V_{ij}-V_{i0}}}{1 + \sum_{k=1}^{J} e^{V_{ik}-V_{i0}}} \end{aligned}$$

$V_{i0} = 0$ と基準化すると、

$$= \frac{e^{V_{ij}}}{1 + \sum_{k=1}^{J} e^{V_{ik}}}.$$

[1]導出の手順は Train 2009 に従っている.

MXL モデル

- IIA 特性を招く第 1 種極値分布に従うパラメータと IIA 特性を引き起こさない分布に従うパラメータに従うパラメータを混合したモデル
- 簡単化のため、$U_{ij} = \mathbf{X}_j'\boldsymbol{\beta}_i + \epsilon_{ij}$ としよう.

$$
\begin{aligned}
U_{ij} &= \mathbf{X}_j'\boldsymbol{\beta}_i + \epsilon_{ij} \\
&= \sum_{k=1}^{K} x_{j,k}\beta_{i,k} + \epsilon_{ij} \\
&= \sum_{k=1}^{K} x_{j,k}(\bar{\beta}_k + \sigma_k \nu_{i,k}) + \epsilon_{ij} \\
&= \underbrace{\sum_{k=1}^{K} x_{j,k}\bar{\beta}_k}_{V_j} + \underbrace{\sum_{k=1}^{K} x_{j,k}\sigma_k \nu_{i,k}}_{z_{j,k}\nu_{i,k}} + \epsilon_{ij}
\end{aligned}
$$

- 誤差項はプロビット分布に従う項と第 1 種極値分布に従う項に分解