離散選択モデル Part 1: フレームワーク

講師: 遠山祐太

最終更新: 2025-01-06

イントロダクション

イントロダクション

- テーマ：離散選択モデルによる構造推定の入門
  - 二項選択モデルの拡張
- 内容：
  - パート1: 離散選択モデルの導入
  - パート2: Rの`mlogit`パッケージを使った実践

注意点など

- スライドタイトルについて：
  - 「【R分析】」：講義パートではサラッとカバー。後ほど説明。
  - 「【参考】」：計量経済学の知識が必要な点。ひとまず割愛してOK

- 補足資料について
  - 補足資料1:理論編 --- 理論の詳細について解説
  - 補足資料2:フルスクラッチでモデル推定 --- Rのパッケージを利用しない離散選択モデルの推定を紹介

参考文献

- RAブートキャンプのトピックレクチャー資料 https://github.com/yutatoyama/discrete_choice_RA_bootcamp
  - 基本的に同じ内容
- 上武・遠山・若森・渡辺 「実証ビジネスエコノミクス」 連載第１回
  - 本講義の内容は連載第１回を書籍向けに拡充したものとなる。
  -
- Croissant "Estimation of Random Utility Models in R: The mlogit Package" *Journal of Statistical Software*
- Train "Discrete Choice Methods with Simulation"
  - 特にChapter 2,3, and 6
- Adams, Clarke, and Quinn "Microeconometrics and Matlab"
- ダハナ・勝又「Rによるマーケティング・データ分析」第５章

構造推定アプローチとは？

構造推定(Structural Estimation)とは？

- 構造推定: 経済主体の意思決定モデルを活用した実証分析のアプローチ
- 手順
    - Step 1: 分析したい経済現象に関する経済モデルを構築する。
    - Step 2: 経済モデル内のパラメタを、データを用いて推定する。
    - Step 3: 推定したモデルを用いて、**反実仮想 (counterfactual)**シミュレーションを行う。

構造推定の例：価格付け

- 問：企業が自社製品の価格付けを検討している。どうすればよいか？
- Step 1：以下の利潤最大化問題を考えよう。

$$\max_p \ (p - mc)D(p)$$

  - p：価格、 mc: 限界費用(＝単位費用), $D(p)$: 需要関数
- Step 2: 需要関数 $D(p)$ と 限界費用 mc を知りたい。
  - 線形な需要関数を考えて、(操作変数法などで)推定する

$$Q_t = \alpha_0 + \alpha_1 \cdot P_t + \beta \cdot income_t + \epsilon_t$$

  - 限界費用は(1) 企業の内部情報、(2) 計量経済学的に推定する (詳細は割愛)
- Step 3: 利潤を最大にするような価格を探す！（数値計算）

価格付け分析の実例

- 例１：Ziprecruiter.com の価格付け実験
  - 米国のオンライン求人サイトが、求人企業の月会費の改定を検討。
  - シカゴ大学の研究者と共同して需要測定のためのA/Bテストを実行。
  - 分析から、価格の最適化により収入を約６０%改善できることが判明。
  - 実際に、価格を９９ドルから２４９ドルへ改定。
- 例２：バッファローのnasneの価格付け
  - ネットワークレコーダーの新製品に関する価格付けの検討。
  - UTEconが、販売予測モデルの構築に携わる。
  - 販売価格29800円が「スイートスポット」にはまり、売れ行きは好調だった。
- 両事例の詳細についてはこちらを参照。

構造推定の強み・弱み

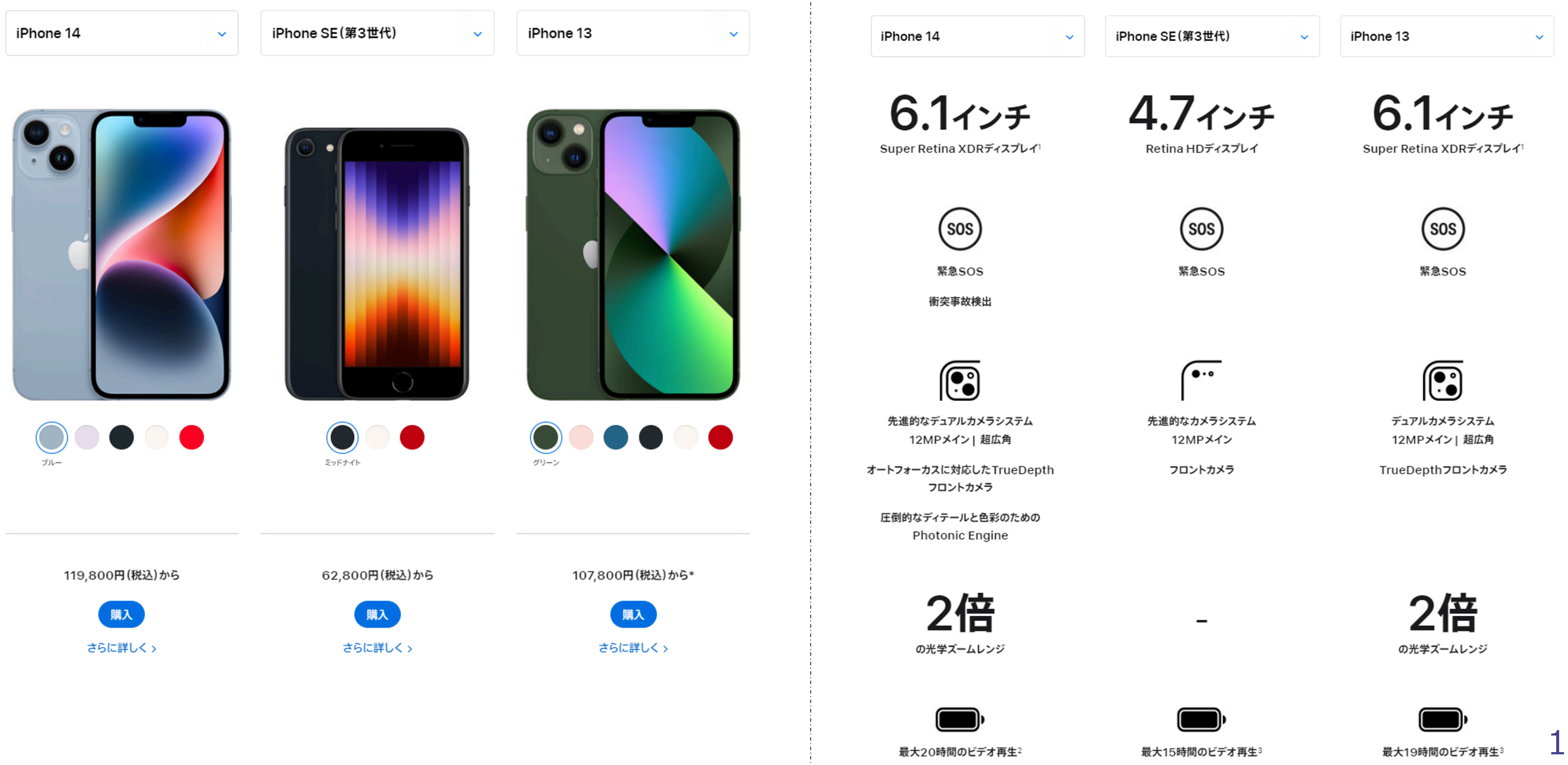
- 強み：経済モデルを用いたシミュレーションができる。
  - 例：ある政策の代わりに別の政策が行われたらどうなっていたか？
  - 例：意思決定におけるあるメカニズム・チャネルを捨象した場合、どうなるか？
  - 因果効果・政策効果の推定の一手段
  - 経済的に関心があるアウトカム（特に経済厚生）を見ることができる。
- 弱み：経済モデルに強く依拠している。
  - 仮定が強い。
  - 分析が複雑になることが多い（プログラミングなど）
- 詳しくは「補足資料1: 理論編」を参照。

今回の例：離散選択モデル

- 本レクチャーでは、**離散選択モデル**を題材として構造推定を紹介する。
- 理由１：応用範囲が非常に広い
  - マーケティング：どの財を購入するか？
  - 交通：通勤経路の選択（電車・バス・自家用車）
  - 労働・教育：どの学校に出願するか？ いつ引退するか？
  - 政治経済学：誰に投票するか？
  - 産業組織論：市場に参入するか否か（ゲーム理論的モデル）
- 理由２：モデル・計量分析・プログラミングともに比較的シンプル
- 理由３：私が一番好きな経済モデルだから。

離散選択モデル

例：あなたはどのiPhoneを買いますか？



離散選択モデル (Discrete Choice Model)

- 離散選択モデル：「**数多くの選択肢の中からどれを選ぶか**」を表現する数理モデル
- 意思決定者は**複数(有限個)の選択肢**に直面している。
  - iPhoneの例：iPhone 13, 14, SE(3世代), 13 Pro, などなど
- それぞれの選択肢(例：製品)には、付随する**特徴**がある
  - iPhoneの例：価格、ディスプレイサイズ、カメラの画素数、バッテリー持続時間、などなど
- 意思決定者は「どの特徴を重視するか」という**好み**をもっている。
  - 例：値段は気にしない（親が払ってくれるから）、大きい画面が良い、などなど
- 選択肢の特徴を踏まえて、**自分の満足度(効用)が最も高くなる選択**をする

セットアップ

- 消費者 $i = 1, \cdots, N$
- 製品・選択肢 $j \in \mathbf{J} = \{1, \ldots, J\}$
  - $\mathbf{J}$ を選択肢集合(choice set)と呼ぶ。
- 消費者 i が製品 j から得られる効用

$$u_{ij} = \alpha p_j + \beta X_j + \epsilon_{ij}$$

  - p_j: 価格
  - X_j: 製品の属性・特徴のベクトル。
  - (α, β): 選好パラメタ
  - ϵ_{ij}: 個人かつ製品レベルの**ランダムな**選好ショック。価格や製品属性で捉えられない要素、言わば誤差項として考える。

簡単な例１：iPhoneの購入

- 以下の効用関数を考えよう

$$u_{ij} = -2p_{jt} + 3 \times (\text{画面サイズ}) + 0.5 \times (\text{バッテリー時間})$$

- 以下の三機種を比較しよう

モデル	価格(万円)	画面サイズ(インチ)	バッテリー(時間)	効用
iPhone 14	11.98	6.1	20	4.34
iPhone SE	6.28	4.7	15	**9.04**
iPhone 13	10.78	6.1	19	6.24

- **iPhone SEを購入**する！！

簡単な例２：もし価格をあまり気にしないと？

- 価格の係数が小さい（＝価格をあまり気にしない）

$$u_{ij} = -1p_{jt} + 3 \times (\text{画面サイズ}) + 0.5 \times (\text{バッテリー時間})$$

- 以下の三機種を比較しよう

モデル	価格(万円)	画面サイズ(インチ)	バッテリー(時間)	効用
iPhone 14	11.98	6.1	20	16.32
iPhone SE	6.28	4.7	15	15.32
iPhone 13	10.78	6.1	19	**17.02**

- **iPhone 13を購入**する！！
- ポイント：人々の「好み」によって、選択は異なってくる！！

離散選択問題の定式化

- 以下のようにノーテーションを定義

$$u_{ij} = V_j + \epsilon_{ij}$$

  - $V_j = \alpha p_j + \beta X_j$.

- 消費者 i は最も高い効用が得られる製品・選択肢 j を一つ選ぶ。

$$d_i = \arg \max_{j \in \{1, \cdots, J\}} u_{ij}$$

離散選択問題から選択確率の導出

- 消費者は効用の要素を把握した上で選択 -> 決定論的選択
- しかしながら、分析者には選好ショック $\{\epsilon_{ij}\}_{j=1}^{J}$ は観察できない。(いわゆる誤差項)
- モデルの予測として、以下の**選択確率** を考える。

$$P(d_i = j) = \Pr\left(\{\epsilon_{ij}\}_{j=1}^{J} : V_j + \epsilon_{ij} \geq V_k + \epsilon_{ik}\ \forall k \neq j\right)$$

  - 解釈：選択肢 j から得られる効用が最も大きくなる確率

- この選択確率は $\{\epsilon_{ij}\}_{j=1}^{J}$ に関する分布の仮定による。
  - 例：ロジットモデル、プロビットモデル

多項ロジットモデル

- $\epsilon_{i,j}$ が i.i.d. の 第一種極値分布(type I extreme value distribution) に従うと仮定。

$$F(x) = exp(-exp(-x))$$

  - 分散は $\pi^2/6$.

- この分布の元で、選択確率は

$$\Pr\left(d_i = j | \{p_j, X_j\}_{j=1}^{J}\right) = \frac{\exp(\alpha p_j + \beta X_j)}{\sum_{k=1}^{J} \exp(\alpha p_k + \beta X_k)}$$

  - **ロジット確率**とも呼ばれる。
  - 導出はTrainのChapter 3を参照。

モデルがあると何が嬉しいか？

- ロジットモデル再掲

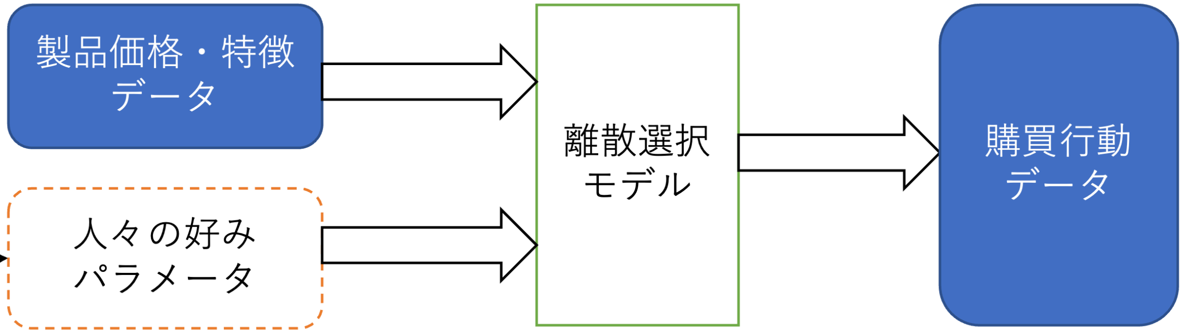
$$\Pr\left(d_i = j | \{p_j, X_j\}_{j=1}^{J}\right) = \frac{\exp(\alpha p_j + \beta X_j)}{\sum_{k=1}^{J} \exp(\alpha p_k + \beta X_k)}$$

- 例１：製品の価格 p_j を下げたら需要がどう変化するか？
  - iPhone SEの値下げ -> SEの需要増加
  - 同時に、iPhone 13の需要は減るかも（製品の代替）
- 例２：新製品を導入したらどうなるか？
  - 例えば、iPhone Max という $J+1$ 番目の新製品 $\{p_{J+1}, X_{J+1}\}$
  - 新しい製品の需要予測、そして既存の製品への影響を予測できる。

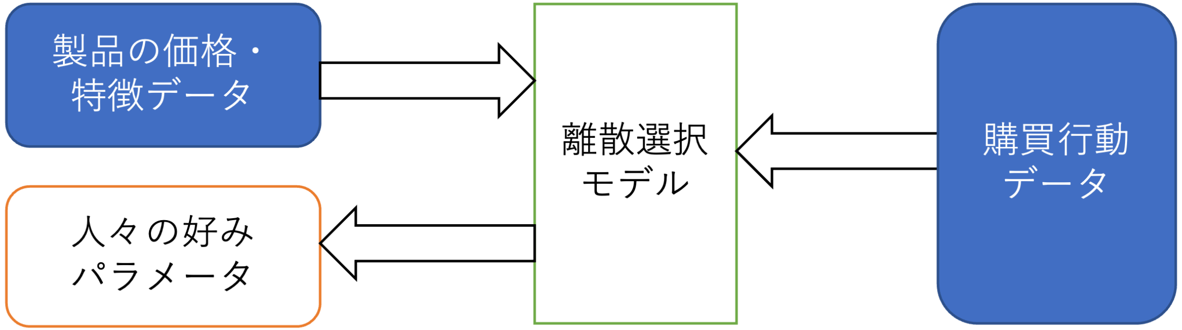
モデルからデータへ：

- モデル

データで
見えないもの！！



- データ分析：



- 「モデルの予測」と「実際の購買行動」が近くなるような「パラメタ」を推定する！！

最尤法 (maximum likelihood) による推定

- データ: 各個人 i について $\{X_j, p_j, d_{ij}\}_{j=1}^{J}$
  - $d_{ij} = 1$ 消費者 i が 製品 j を選んだとき 1，それ以外は0
- **尤度関数(likelihood)**: ある実現値が発生するような確率をパラメタの関数と表したもの。
- 多項ロジットモデルの尤度関数 (パラメタ $\theta = (\alpha, \beta)$)

$$L(\theta) = \prod_{i=1}^{N} \left[\prod_{j=1}^{J} (Pr(d_i = j|\theta))^{d_{ij}} \right]$$

- 尤度関数は、「手元にあるデータが、モデルによって生成される確率」と解釈される。
- この確率を最大にするようなパラメタを「良い推定量」とする。
  - 一定の仮定のもとで一致性・漸近正規性・効率性などの理論的性質が担保されている。

対数尤度関数の最大化

- 最尤法では、以下の対数尤度を用いる。

$$\log L(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log Pr(d_i = j|\theta)$$

- 対数尤度関数を最大化するようなパラメタ θ が最尤推定量となる。
- ポイント：最大化するパラメタは、基本的に数値計算・数値最適化で求める。
  - 非常にシンプルなモデル（例：線形回帰モデル）では解析的に得られる。