

応用計量経済学 (2): 離散選択モデルの特性

Tetsuya Shimane

Apr. 17, 2015

選択集合

選択集合とは

選択集合の例: スマートフォン OS の選択

互いに背反な選択肢

網羅的な選択肢

有限な選択肢数：回帰モデルか離散選択モデルか

有限な選択肢数：回帰モデルか離散選択モデルか

選択確率の導出

様々な離散選択モデル

選択モデルの識別

集計と予測

選択集合

選択集合とは

- ✓ 離散選択モデル (Discrete choice models): 意思決定者が（直面した）選択肢から選び取る行動を記述する.
(Choice set (選択集合): alternatives(選択肢) の集合)
- ✓ 選択集合に求められる要件
 1. 互いに背反 $\Rightarrow$ 一つだけを選ぶ
 2. 網羅的 $\Rightarrow$ 必ずいずれかを選ぶ
 3. 選択肢の数は有限でなければならない1, 2 については多くの場合, 排他的かつ網羅的な選択肢をとることができる.
- ✗ 選択肢 A, B が, $A \cap B \neq \emptyset$: A, B が互いに背反ではない.
 $\Rightarrow$ 選択肢: $A \cap B, A \cap B^c, A^c \cap B$ は互いに排反 (X^c : X の補集合)
- ✗ 選択肢 A, B で, $A^c \cap B^c$ が存在する: 網羅的ではない.
 $\Rightarrow$ 選択肢: $A, B, A^c \cap B^c$ は網羅的

選択集合の例: スマートフォン OS の選択

- ✓ 購入するスマートフォンの OS を選択
- ✓ 選択肢は {Android}, {iOS} の二つとして良いか？
 - ✗ 両方とも導入: 互いに背反ではない
⇒ {Android}, {iOS}, {Android& iOS}
 - ✗ その他の OS の利用: 網羅的ではない
⇒ {Android}, {iOS}, {others}
 - ✗ {Android}, {iOS}, {others}, {Android& iOS}, {Android& others}, {iOS& others}, {Android, iOS, others} とすれば, 互いに排反かつ網羅的
- ✓ 他に方法はないのか？

互いに背反な選択肢

互いに背反な選択肢を作るための対処法：

1. 可能な選択肢の組み合わせをすべてリストする
2. “主要な” 選択を定義する

何れの方法をとるべきかは分析の目的，使用できるデータで判断する．

1. “主要な” を決める困難がない．
複数スマートフォンを購入するものについて分析可能
組み合わせごとの情報が必要
2. データの要求が少ない
推定，予測は選択肢の数が少ないほど容易

網羅的な選択肢を作るための対処法：

1. その他の選択肢を導入する
2. 選択の状況を再定義する: 最も代表的な OS を導入したスマートフォンにおける選択とする.

これも、分析の目的や利用できるデータで判断する.

1. その他の OS(Windows Phone, Tizen, etc.) の利用についても調べられる
その他に関して意味の有るデータを用意する必要がある.(その他に複数の選択肢が含まれると難しくなる)
2. そのようなデータの必要がから解放される.

有限な選択肢数：回帰モデルか離散選択モデルか

- ✓ 回帰モデルと離散選択モデルの違いは,
 - ✗ 回帰モデル: 可能な結果の数が無限 $\Rightarrow$ Choice of “how much”
 - ✗ 離散選択モデル: 可能な結果の数が有限 $\Rightarrow$ Choice of “which”

$\rightarrow$ 正確な理解ではない.
- ✓ 例: 家計の保有自動車数の選択
家計が保有する自動車の数をどのように決めているかを調べたい.
このとき保有する自動車の数は, 選択肢 = $\{0, 1, 2, \dots\}$ であり, 理論的には有限とはいえない. (もしくは計算が実行可能な選択肢の数を上回る)
 - ✗ この選択集合は互いに背反かつ網羅的な有限この選択肢を含んでいる. 選択肢 = $\{0, 1, 2 \text{ or more}\}$
 - ✗ 離散選択モデルによる分析に適當 (なぜ?)
 - ✗ 預金残高でも適用できるか?

有限な選択肢数：回帰モデルか離散選択モデルか

- ✓ 回帰モデルと離散選択モデルどちらを使うか
 - ✗ 回帰モデルはより自然かつ簡単
 - ✗ (非線形価格等の) 理由がある場合には、離散選択モデルを採用
 - ✗ 研究の目的と離散選択モデルの可能性を考慮して判断する必要がある。

選択集合

選択確率の導出

確率効用モデル

選択確率の導出

選択確率の導出

観察されない要素 ε_n

選択確率の意味

選択確率の意味

観察されない要素 ε_n
の分布とは何か

様々な離散選択モデル

選択モデルの識別

集計と予測

選択確率の導出

離散選択モデルは意思決定者の効用最大化行動の結果のもとで導出される。(Random Utility Model: 確率効用モデル)

モデル

- ✓ 意思決定者 n が J 個の選択肢からの選択に直面しているとする.
選択肢: $j = 1, \dots, J$, 効用: $U_{nj}, j = 1, \dots, J$
- ✓ U_{nj} は, 意思決定者には既知, しかし分析者には未知.
- ✓ n は選択肢 i を選択 iif $U_{ni} > U_{nj}, \forall j \neq i$
- ✓ 分析者からは
観察できないもの: U_{nj}
観察できるもの: x_{nj} $\forall j$: 選択肢の属性, S_n : 意思決定者の属性
- ✓ U_{nj} について x_{nj}, S_n の関数を特定化する: representative utility
 $V_{nj} := V(x_{nj}, S_n) \Leftarrow$ この関数の未知パラメータを推定する.

選択確率の導出

- ✓ 分析者は効用を観察できないので $U_{nj} \neq V_{nj}$. 効用は次のように分解される.
 $U_{nj} = V_{nj} + \varepsilon_{nj}$ ε_{nj} : V_{nj} に含まれない効用に影響を与える要素
- ✓ $\varepsilon_{nj} \forall j$ は未知なので, ランダムとして扱う.
 $\varepsilon_n = (\varepsilon_{n1}, \dots, \varepsilon_{nJ})$ の同時分布を同時密度関数 $f(\varepsilon_n)$ で表す.
 $\Rightarrow$ 分析者は意思決定者の選択について確率的に述べることができる.

$$\begin{aligned} P_{ni} &= \Pr(U_{ni} > U_{nj} \quad \forall j \neq i) \\ &= \Pr(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \quad \forall j \neq i) \\ &= \Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \quad \forall j \neq i) \end{aligned}$$

選択確率 P_{ni} とは全てのランダム項 $\varepsilon_{nj} - \varepsilon_{ni}$ が $V_{ni} - V_{nj}$ より小さいの確率.

- ✓ 分布 $f(\varepsilon_n)$ を用いて ($I(\cdot)$ は指示関数)

$$\begin{aligned} P_{ni} &= \Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \quad \forall j \neq i) \\ &= \int_{\varepsilon} I(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \quad \forall j \neq i) f(\varepsilon_n) d\varepsilon_n \end{aligned}$$

観察されない効用の部分の分布 $f(\varepsilon_n)$ に関する重積分.

- ✓ $f(\varepsilon)$ は特定化次第.
- ✗ Closed form で得られるのは:
 - ✓ Logit: f は iid 極値分布 (EV).
 - ✓ Nested logit: 一般化極値分布 (GEV) の一種.
- ✗ Closed form で得られない : Simulation の活用
 - ✓ Probit: 多次元正規分布
 - ✓ Mixed logit: 一部は分析者の特定化により, 一部は iid 極値分布

選択確率の意味

- ✓ 通勤手段の選択: 自家用車 (c) or バス (b)
- ✓ 観察される要素は, 所要時間 (T), 金銭的費用 (M). 他は不明 (観察できない).
- ✓ representative utility を次のように特定化した.

$$V_c = \alpha T_c + \beta M_c$$

$$V_b = \alpha T_b + \beta M_b$$

α, β は推定されている (or 既知) とする.

- ✓ 与えられた α, β で $V_c = 4, V_b = 3$ となったとする.

選択確率の意味

- ✓ 自家用車を必ず選ぶことを意味しない! 観察されない要素がある.
- ✓ 観察されない要素が1より大きければバスを選ぶ.
(観察されない要素が $V_c - V_b$ より大きければ) $\Rightarrow \varepsilon_b - \varepsilon_c > 1$
- ✓ 自家用車を選ぶのは $\varepsilon_b - \varepsilon_c < 1$ の場合.

$$P_c = \Pr(\varepsilon_b - \varepsilon_c < V_c - V_b)$$

$$P_b = \Pr(\varepsilon_b - \varepsilon_c > V_c - V_b) = \Pr(\varepsilon_c - \varepsilon_b < V_b - V_c)$$

観察されない要素 ε_n の分布とは何か

特定化を決定づけている ε_n とは何か? 選択確率とは何か?

- ✓ 同じ観察された効用の部分を持つ人々の母集団における効用の観察されない部分の分布. (頻度主義的)
- ✓ 個人の効用の観察できない部分についての分析者の考える主観的な確率. (主観的確率)
- ✓ 個人の衝動的な行動により, 合理的とはいえない他の選択肢をとってしまう確率.

選択集合

選択確率の導出

様々な離散選択モデル

様々な離散選択モデル:
Logit model

様々な離散選択モデル:
GEV model

様々な離散選択モデル:
Probit model

様々な離散選択モデル:
Mixed logit model

選択モデルの識別

集計と予測

様々な離散選択モデル

様々な離散選択モデル: Logit model

異なる選択モデル $\Leftrightarrow$ 観測されない要素の分布 $f(\varepsilon_n)$ の異なる特定化

- ✓ どんな分布を仮定しているのか？
- ✓ これらの仮定を導入するモチベーションは何か？

Logit model

- ✓ ε_{ni} : iid extreme value for all i
 $f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}}$, $F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}}$
- ✓ 観測されない要素は選択肢間で相関しない. また同一の分布 $\Rightarrow$ すべての選択肢は同一の分散.
- ✓ 制約的だが便利な手法
- ✓ 系時的に一続きの選択に適用した場合, それぞれの選択は独立と仮定されてしまう.

GEV model

- ✓ 独立性の仮定を回避するために, ε_n に一般化極値分布 (GEV) を採用.
- ✓ GEV の採用により, 観測されない要素の選択肢間の相関を許容するモデルとなっている. (例: Nested logit)
- ✓ GEV model は選択確率を closed form で記述できる.

様々な離散選択モデル: Probit model

Probit model

- ✓ 観測されない要素が同時正規分布に従うとする.
 $\varepsilon'_n = (\varepsilon_{n1}, \dots, \varepsilon_{nJ}) \sim N(0, \Omega).$
- ✓ Ω について, どんなパターンの相関も異分散も扱うことができる.
- ✓ 系時的な一続きの選択の扱いが容易: 観察されない要因は, 選択肢の間と同様に, 時点間の同時正規分布として仮定される.
- ✓ Probit model の長所: 選択肢間, 時点間の相関を柔軟に取り扱える.
- ✓ Probit model の限界: 観察されない要因は正規分布に従わないかもしれない.

様々な離散選択モデル: Mixed logit model

Mixed logit model

- ✓ 観測されない要素を (分析者が特定化した) 任意の分布に従うとすることができる.
- ✓ (観察されない要素) = (相関, 異分散をもたらす要素) + (iid EV)

それぞれのモデルの導出とモチベーションを理解することによって, 分析をする際に, 状況と目的に合わせたモデルを特定化することができるようになる.

選択集合

選択確率の導出

様々な離散選択モデル

選択モデルの識別

選択モデルの識別:

Identification of choice models

効用の差だけが重要

選択肢 j 固有の定数項

例：通勤手段の選択 自動車-バス

意思決定者 n 固有の変数

独立な誤差項の数

効用の全体の規模は問題ではない

誤差の基準化：

$\varepsilon_{nj} \sim iid$ の場合

誤差の基準化：

$\text{Var}(\varepsilon_{nj}) \neq \sigma^2$

誤差の基準化：

$\text{Cov}(\varepsilon_{nj}, \varepsilon_{nj'}) \neq 0$

集計と予測

選択モデルの識別

選択モデルの識別: Identification of choice models

離散選択モデルを使うことで、
何が分かり（識別され）、何が分からない（識別されない）のか。

離散選択モデルの識別において注目すべき点

- ✓ 効用については差だけ: Only differences in utility matter
- ✓ 効用の規模は任意: The scale of utility is arbitrary

効用の差だけが重要

- ✓ 効用の絶対的水準は意思決定者の行動にも研究者のモデルにも関係ない.
例: (1) U_{nj} , (2) $U_{nj} + k$ (k : Constant) ← 何れの選択肢にも同じオマケをつける
or キャッシュバック
⇒ (1), (2) で同じ選択をする.
- ✓ 意思決定者: $U_{nj} \forall j$ の場合も, 任意の一定の k について $U_{nj} + k \forall j$ の場合も同じ
選択肢を選ぶ. (最も効用の高い選択肢は同じ)
- ✓ 研究者: 選択確率 $P_{ni} = \Pr(U_{ni} > U_{nj} \forall j \neq i) = \Pr(U_{ni} - U_{nj} > 0 \forall j \neq i)$
効用の差 だけ に依存する. 絶対的水準ではない.
分解した場合でもこれは変わらない. $P_{ni} = \Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \forall j \neq i)$
- ✓ モデルの同定や特定化をすることから見ると:
推定できる (識別できる) パラメータとは, 選択肢の間の差 (違い) を捉えるもの
だけである.

選択肢 j 固有の定数項

- ✓ 効用の観察できる部分をパラメータに関して、定数項を伴った、線型で特定化する.

$$V_{nj} = x'_{nj}\beta + k_j \quad \forall j,$$

ただし, x_{nj} : 意思決定者 n が直面する選択肢 j に関する変数のベクトル

β : 変数 x_{nj} の係数ベクトル

k_j : 選択肢 j に固有の定数

$\Leftrightarrow$ モデルに含まれない全ての要素の平均的な効果

- ✓ IF k_j がふくまれる, THEN $\Rightarrow E(\varepsilon_{nj}) = 0$
 $\therefore$ IF $U_{nj} = x'_{nj}\beta + \varepsilon^*_{nj}$ with $E(\varepsilon^*_{nj}) = k_j \neq 0$,
THEN $U_{nj} = x'_{nj}\beta + k_j + \varepsilon_{nj}$ with $E(\varepsilon_{nj}) = 0$
- ✓ 効用の差だけが問題なので, 選択肢固有の定数項の差だけが問題とされ, 絶対的水準は関係ない!

例：通勤手段の選択 自動車-バス

自動車 (c) で通勤するか, バス (b) で通勤するか. それぞれの場合に得られる効用を U_c, U_b とする. 所要時間 T_j , 費用 M_j で, representative utility が構成されるとすると.

$$\begin{cases} U_c = \alpha T_c + \beta M_c + k_c^0 + \varepsilon_c, \\ U_b = \alpha T_b + \beta M_b + k_b^0 + \varepsilon_b, \end{cases} \quad \text{with } k_b^0 - k_c^0 = d$$
$$\begin{cases} U_c = \alpha T_c + \beta M_c + k_c^1 + \varepsilon_c, \\ U_b = \alpha T_b + \beta M_b + k_b^1 + \varepsilon_b, \end{cases} \quad \text{with } k_b^1 - k_c^1 = d$$

2 つの d が同一 $\Rightarrow$ 識別可能

$$\Leftrightarrow \begin{cases} U_c = \alpha T_c + \beta M_c + \varepsilon_c, \\ U_b = \alpha T_b + \beta M_b + k_b + \varepsilon_b, \end{cases} \quad \text{with } k_b = d \quad k_c = 0 \text{ で規準化}$$

$J-1$ 個の選択肢固有定数項をモデルに加えることができる.

意思決定者 n 固有の変数

意思決定者 n 固有の変数 (Sociodemographic variables)

- ✓ 意思決定者の属性は選択肢の間で変化しない.
選択肢間で効用に差をもたらすような特定化をした場合にだけモデルに加えることができる.
- ✓ 例：通勤手段の選択 自動車 vs バス
意思決定者の属性 Y : 所得

$$U_c = \alpha T_c + \beta M_c + \theta_c^0 Y + \varepsilon_c,$$

$$U_b = \alpha T_b + \beta M_b + \theta_b^0 Y + k_b + \varepsilon_b$$

θ_c^0, θ_b^0 : 自家用車かバスに乗ることの効用に所得の違いが与える影響

$\theta_c^0 > 0, \theta_b^0 > 0$ かつ $\theta_c^0 \neq \theta_b^0$ が期待される.

θ_c^0 と θ_b^0 は推定されない, その差だけが推定される.



$$U_c = \alpha T_c + \beta M_c + \varepsilon_c,$$

$$U_b = \alpha T_b + \beta M_b + \theta_b Y + k_b + \varepsilon_b$$

ここで, $\theta_b = \theta_b^0 - \theta_c^0$: 自家用車に比べてバスに乗ったことの効用に所得の与える影響の差.

θ_b は正も負もとりうる.



別の特定化

$$U_c = \alpha T_c + \beta M_c / Y + \varepsilon_c,$$

$$U_b = \alpha T_b + \beta M_b / Y + \theta_b Y + k_b + \varepsilon_b,$$

独立な誤差項の数

- ✓ 効用の差だけが問題であることを考慮すると、次元を減らすことができる.

$$P_{ni} = \int_{\varepsilon} I(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \quad \forall j \neq i) f(\varepsilon_n) d\varepsilon_n$$

ε は誤差の J 次元ベクトル

$$\begin{aligned} P_{ni} &= \Pr(U_{ni} > U_{nj} \quad \forall j \neq i) \\ &= \Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \quad \forall j \neq i) \\ &= \Pr(\tilde{\varepsilon}_{nji} < V_{ni} - V_{nj} \quad \forall j \neq i) \\ &= \int I(\tilde{\varepsilon}_{nji} < V_{ni} - V_{nj} \quad \forall j \neq i) g(\tilde{\varepsilon}_{nji}) d\tilde{\varepsilon}_{ni} \end{aligned}$$

ただし $\tilde{\varepsilon}_{nji} = \varepsilon_{nj} - \varepsilon_{ni}$,

$\tilde{\varepsilon}_{ni} = (\tilde{\varepsilon}_{n1i}, \dots, \tilde{\varepsilon}_{n(i-1)i}, \tilde{\varepsilon}_{n(i+1)i}, \dots, \tilde{\varepsilon}_{nJi}) \quad (J-1) \text{次元ベクトル}$

- ✓ 選択確率は $g(\tilde{\varepsilon}_{nk})$ だけに従うように表現され、分布 $f(\varepsilon_n)$ の 1 つの次元は識別されず、研究者により基準化されなければならない。

効用の全体の規模は問題ではない

- ✓ 効用がどのような尺度 (スケール) をとっても、最も効用の高い選択肢は変わらない.

$$\begin{aligned} U_{nj}^0 &= V_{nj} + \varepsilon_{nj} \quad \forall j \\ \Leftrightarrow U_{nj}^1 &= \lambda V_{nj} + \lambda \varepsilon_{nj} \quad \forall j \quad \text{for any } \lambda > 0 \end{aligned}$$

研究者が効用のスケールを基準化しなければならない.

$\Rightarrow$ 誤差項の分散を基準化 $\because \lambda U_{nj} \Rightarrow \varepsilon_{nj}$ は λ^2 で変わる. $\because \text{Var}(\lambda \varepsilon_{nj}) = \lambda^2 \text{Var}(\varepsilon_{nj})$

- ✓ 基準化
 - ✗ 独立同分布 (iid) 誤差
 - ✗ 異分散の誤差
 - ✗ 相関のある誤差

誤差の基準化： $\varepsilon_{nj} \sim iid$ の場合

- ✓ 全ての誤差項が同じ分散を持つので、直接基準化する。
基準化前のモデル

$$U_{nj}^0 = x'_{nj}\beta + \varepsilon_{nj}^0 \quad \text{where } \text{Var}(\varepsilon_{nj}^0) = \sigma^2.$$

- ✓ 誤差項の分散を 1 として基準化した場合:

$$U_{nj}^1 = x'_{nj}(\beta/\sigma) + \varepsilon_{nj}^1 \quad \text{with } \text{Var}(\varepsilon_{nj}^1) = 1$$

β/σ : 観測されない要素 ε_{nj} の標準偏差と比べた、観測された変数 x_{nj} の効果。
 $\Rightarrow$ 観測されない要素の振る舞いの大きさ次第で、係数 β/σ は変化する。

誤差の基準化： $\varepsilon_{nj} \sim iid$ の場合

- ✓ Logit の場合: 誤差の分散は $\pi^2/6 \simeq 1.6$ に基準化される.

$$U_{nj} = x'_{nj}(\beta/\sigma) \sqrt{1.6} + \varepsilon_{nj} \quad \text{with } \text{Var}(\varepsilon) = 1.6$$

係数 $(\beta/\sigma) \sqrt{1.6}$ は依然として観測されない要素 ε_{nj} の分散次第である.
係数は分散=1 と基準化した場合に比べ $\sqrt{1.6}$ 倍となる.

- ✓ Probit と Logit の推定された係数の比較
- ✓ 異なるデータで同じモデルの係数の比較

誤差の基準化： $\text{Var}(\varepsilon_{nj}) \neq \sigma^2$

- ✓ ある一つのセグメントの分散によって、効用のスケールを基準化し、そのセグメントと比べた各セグメントの分散を推定する.
- ✓ Chicago-Boston データセット：2つの都市 (Chicago と Boston) で、調査をしたデータをまとめてモデルに適用する.
しかし、両都市で観察されない要素が異なると、誤差項の分散は異なってくる.

$$U_{nj} = \alpha T_{nj} + \beta M_{nj} + \varepsilon_{nj}^B \quad \forall n \quad \text{in Boston,}$$

$$U_{nj} = \alpha T_{nj} + \beta M_{nj} + \varepsilon_{nj}^C \quad \forall n \quad \text{in Chicago,}$$

$$\text{where } \text{Var}(\varepsilon_{nj}^B) \neq \text{Var}(\varepsilon_{nj}^C).$$

誤差の基準化 : $\text{Var}(\varepsilon_{nj}) \neq \sigma^2$

- ✓ 両都市の分散の比を $k := \text{Var}(\varepsilon_{nj}^C) / \text{Var}(\varepsilon_{nj}^B)$ とおくと,

$$U_{nj} = \alpha T_{nj} + \beta M_{nj} + \varepsilon_{nj} \quad \forall n \quad \text{in Boston,}$$

$$U_{nj} = (\alpha \sqrt{k}) T_{nj} + (\beta \sqrt{k}) M_{nj} + \varepsilon_{nj} \quad \forall n \quad \text{in Chicago.}$$

ここで, 両都市の (重み付けされた) 観測されない要素 ε_{nj} の分散は等しい.

$$\therefore \text{Var}(\varepsilon_{nj}^C / \sqrt{k}) = \text{Var}(\varepsilon_{nj}^B)$$

- ✓ スケールパラメータ k は β, α とともに推定される.
 $\hat{k}$: ボストンと比べたシカゴの観測されない要素の分散
 $\hat{k} = 1.2$ であれば, シカゴの方が観測されない要素の分散が 20% 大きいと解釈される.
- ✓ 誤差項 (観測されない要因) の分散は, 地理的な違い, データセットの違い, 調査した時期の違いなどの要因で異なると考えられる.

誤差の基準化： $\text{Cov}(\varepsilon_{nj}, \varepsilon_{nj'}) \neq 0$

- ✓ 選択肢の間で誤差が相関している場合には、基準化はより複雑。
「効用のスケールの設定」 $\Rightarrow$ 「効用の差のスケールの設定」
- ✓ 誤差が相関している場合、一つを選択肢について誤差の分散を正規化することでは、十分ではない!!

例: 4 選択肢の場合

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \quad j = 1, \dots, 4$$

$$\varepsilon_n = (\varepsilon_{n1}, \dots, \varepsilon_{n4})'$$

$$\varepsilon_n \text{ は平均ゼロ, 共分散 } \Omega = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \cdot & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \cdot & \cdot & \sigma_{33} & \sigma_{34} \\ \cdot & \cdot & \cdot & \sigma_{44} \end{bmatrix}$$

誤差の基準化：Cov($\varepsilon_{nj}, \varepsilon_{nj'}$) $\neq 0$ Difference form

「効用の差だけが問題」なので，ある選択肢との差で各選択肢の効用を表現しても，同値である．

$$\tilde{U}_{nj1} = \tilde{V}_{nj1} + \tilde{\varepsilon}_{nj1} \quad \text{for } j = 2, 3, 4$$

$$\text{ここで } \tilde{U}_{nj1} = U_{nj} - U_{n1}$$

$$\tilde{V}_{nj1} = V_{nj} - V_{n1}$$

$$\tilde{\varepsilon}_{n1} = ((\varepsilon_{n2} - \varepsilon_{n1}), (\varepsilon_{n3} - \varepsilon_{n1}), (\varepsilon_{n4} - \varepsilon_{n1}))'$$

$\tilde{\varepsilon}_{n1}$ の分散共分散は元の誤差 ε_n の分散共分散により決まる．

$$\begin{aligned} \text{Var}(\varepsilon_{\tilde{n}21}) &= \text{Var}(\varepsilon_{n2} - \varepsilon_{n1}) = \text{Var}(\varepsilon_{n1}) + \text{Var}(\varepsilon_{n2}) - 2 \text{Cov}(\varepsilon_{n1}, \varepsilon_{n2}) \\ &= \sigma_{11} + \sigma_{22} - 2\sigma_{12} \end{aligned}$$

$$\begin{aligned} \text{Cov}(\tilde{\varepsilon}_{n21}, \tilde{\varepsilon}_{n31}) &= E(\varepsilon_{n2} - \varepsilon_{n1})(\varepsilon_{n3} - \varepsilon_{n1}) = E(\varepsilon_{n2}\varepsilon_{n3} - \varepsilon_{n2}\varepsilon_{n1} - \varepsilon_{n3}\varepsilon_{n1} + \varepsilon_{n1}\varepsilon_{n1}) \\ &= \sigma_{23} - \sigma_{21} - \sigma_{31} + \sigma_{11} \end{aligned}$$

誤差の基準化： $\text{Cov}(\varepsilon_{nj}, \varepsilon_{nj'}) \neq 0$

- ✓ Ω から, difference form に対応する $\tilde{\Omega}$ を導くと,

$$\tilde{\Omega}_1 = \begin{bmatrix} \sigma_{11} + \sigma_{22} - 2\sigma_{12} & \sigma_{11} + \sigma_{23} - \sigma_{12} - \sigma_{13} & \sigma_{11} + \sigma_{24} - \sigma_{12} - \sigma_{14} \\ \cdot & \sigma_{11} + \sigma_{33} - 2\sigma_{13} & \sigma_{11} + \sigma_{34} - \sigma_{13} - \sigma_{14} \\ \cdot & \cdot & \sigma_{11} - \sigma_{44} - 2\sigma_{14} \end{bmatrix}$$

- ✓ 誤差の差の分散を定めるためには, 元の「誤差」の一つについて分散を定めることでは不十分である.
(例: $\sigma_{11} = k$ としたとき, $k + \sigma_{22} - 2\sigma_{12}$ は不定となる)
- ✓ $\Rightarrow$ 「誤差の差」の一つについて分散をある数に基準化.

誤差の基準化： $\text{Cov}(\varepsilon_{nj}, \varepsilon_{nj'}) \neq 0$

$\text{Var}(\tilde{\varepsilon}_{n21}) = 1$ と基準化.

$$\begin{aligned}\hat{\Omega}_1^* &= \begin{bmatrix} 1 & \frac{\sigma_{11} + \sigma_{23} - \sigma_{12} - \sigma_{13}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} & \frac{\sigma_{11} + \sigma_{24} - \sigma_{12} - \sigma_{14}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \\ \cdot & \frac{\sigma_{11} + \sigma_{33} - 2\sigma_{13}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} & \frac{\sigma_{11} + \sigma_{34} - \sigma_{13} - \sigma_{14}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \\ \cdot & \cdot & \frac{\sigma_{11} - \sigma_{44} - 2\sigma_{14}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \frac{\sigma_{11} + \sigma_{23} - \sigma_{12} - \sigma_{13}}{m} & \frac{\sigma_{11} + \sigma_{24} - \sigma_{12} - \sigma_{14}}{m} \\ \cdot & \frac{\sigma_{11} + \sigma_{33} - 2\sigma_{13}}{m} & \frac{\sigma_{11} + \sigma_{34} - \sigma_{13} - \sigma_{14}}{m} \\ \cdot & \cdot & \frac{\sigma_{11} - \sigma_{44} - 2\sigma_{14}}{m} \end{bmatrix}\end{aligned}$$

ここで $m = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$. このスケールを得るためには、効用は $\sqrt{\sigma_{11} + \sigma_{22} - 2\sigma_{12}}$ で除される.

誤差の基準化： $\text{Cov}(\varepsilon_{nj}, \varepsilon_{nj'}) \neq 0$

- ✓ 共分散行列中の推定されるパラメータの数
 - ✗ Ω は 10 要素
 - ✗ $\tilde{\Omega}_1$ は 6 要素,
 - ✗ $\tilde{\Omega}_1$ の一つのパラメータについて基準化した $\tilde{\Omega}_1^*$ は 5 パラメータのみ!!
- ✓ $\Rightarrow$ “Only differences matter”, “the scale of utility is arbitrary” が分かれば分散共分散パラメータの数は 10 から 5 つにまで減る.
- ✓ $\Rightarrow J$ 選択肢モデルは高々 $J(J-1)/2 - 1$ の共分散パラメータを持つに過ぎない.
(しかし, 共分散行列のパラメータの解釈は困難になる.)
- ✓ 注意: Logit or nested logit の規準化は分布の仮定より自動的になされる. Mixed logit or probit の規準化は自動的になされない $\Rightarrow$ モデルの特定化や解釈の際に基準化の問題に配慮する必要がある.

- [選択集合](#)
- [選択確率の導出](#)
- [様々な離散選択モデル](#)
- [選択モデルの識別](#)
- [集計と予測](#)**
 - 集計
 - 離散選択モデルの集計測度: 平均確率
 - 離散選択モデルの集計測度: 平均反応（限界効果）
 - Saple enumeration
 - Segmentation
- 予測
 - 定数項の再キャリブレーション

集計と予測

- ✓ 集計測度 (aggrigation measure) は
 - ✗ 母集団での平均確率
 - ✗ ある要素の変化に対する反応の平均を吟味したいときに必要となる.
- ✓ 線形モデル $h_n = \alpha + \beta y_n$ の場合は簡単
 - ✗ average outcome: $\alpha + \beta \bar{y}$
 - ✗ average response to one unit change y_n : β
- ✓ 離散選択モデルでは, 平均値をモデルに挿入することは平均確率や平均反応の不偏推定を与えない!
以下図解

離散選択モデルの集計測度：平均確率

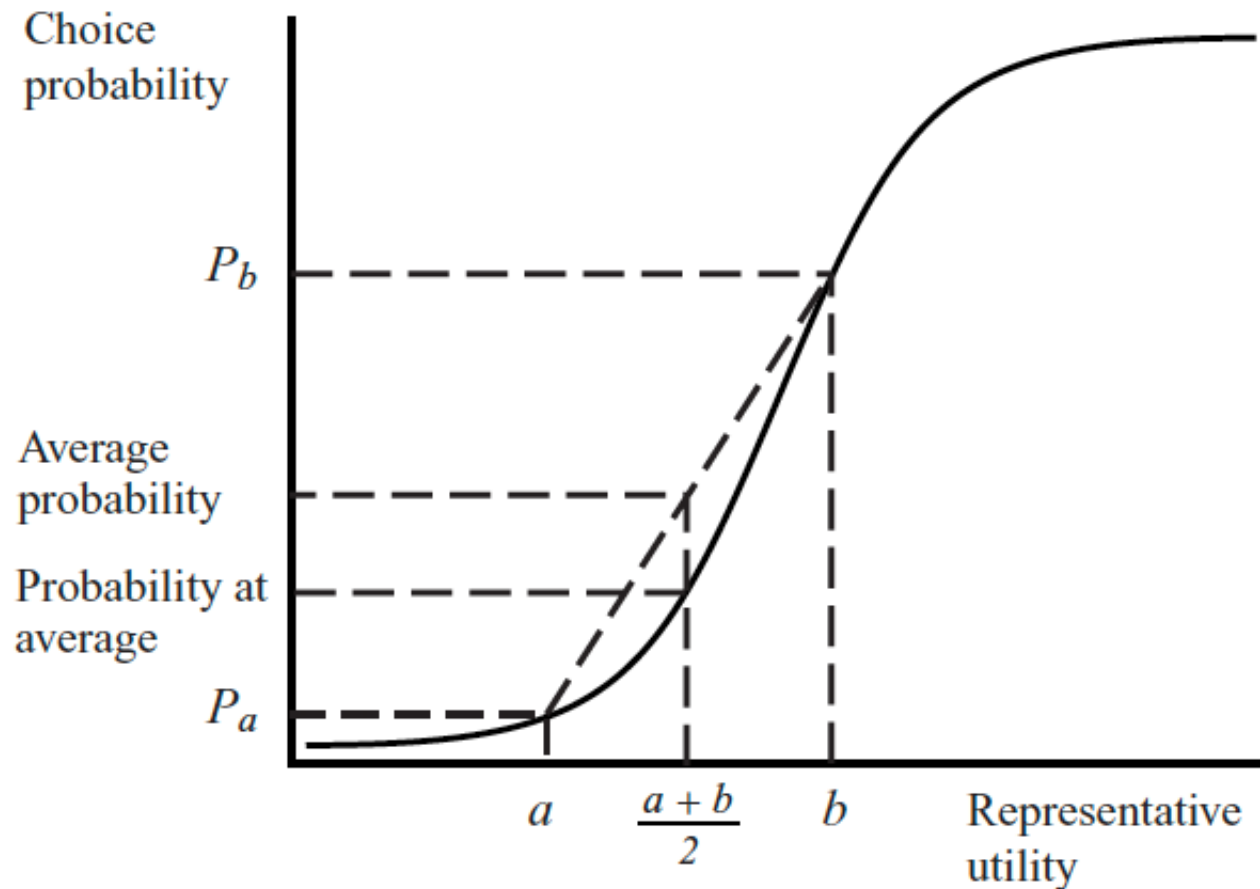


Figure 2.1. Difference between average probability and probability calculated at average representative utility.

離散選択モデルの集計測度: 平均反応 (限界効果)

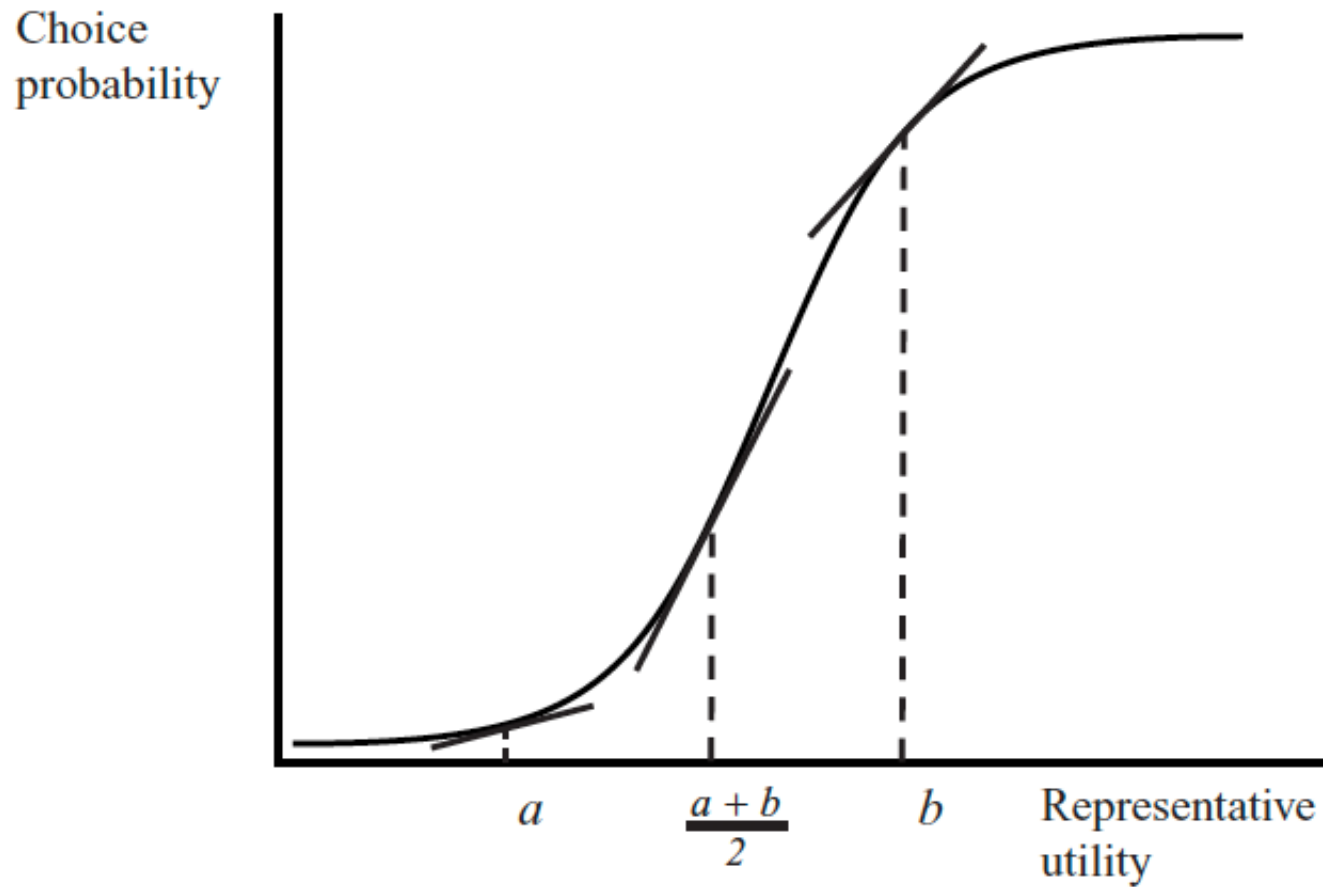


Figure 2.2. Difference between average response and response calculated at average representative utility.

Sample enumeration

- ✓ サンプルの各意思決定者の選択確率を合計もしくは平均する.
- ✓ 母集団における選択肢 i を選ぶ意思決定者の合計数の一致推定値 $\hat{N}_i$ は, この確率の重み付け平均である.

$$\hat{N}_i = \sum_n w_n P_{ni}$$

w_n : 意思決定者 n の重み

平均確率 (市場シェア) : $\hat{N}_i / N$

- ✓ Average derivatives, Average elasticities についても同様.

Segmentation

- ✓ 説明変数の数が少ないときかつこれらの変数が少ない数の値のみをとる場合に利用できる.

✓

	$x_2 = 0$	$x_2 = 1$
$x_1 = 0$	P_{00}	P_{01}
$x_1 = 1$	P_{10}	P_{11}

$$\Rightarrow \hat{N}_i = \sum_{s=1}^4 w_s P_{si}$$

P_{si} : segment s で選択肢 i を選ぶ意思決定者の確率

w_s : segment s の意思決定者の数

集計測度に関する議論を適用する.

- ✓ Sample enumeration
- ✓ Segmentation

外生変数 and/or 重みを予測する将来にドローしてきたように修正する.

定数項の再キャリブレーション

(将来) 予測などを行う場合には, 必要に応じて選択肢固有の定数項を改めて収束計算することでより尤もらしい予測を得ることができる.

1. α_j^0 : 推定された選択肢 j 固有の定数項,
 S_j : 基準年における, 予測領域での選択肢 j を選んだ意思決定者のシェア,
2. $S_j^0 \quad \forall j$: 予測領域での選択肢 j を選ぶ意思決定者の予測シェア,

$$\alpha_j^1 = \alpha_j^0 + \ln(S_j / \hat{S}_j^0).$$

3. このプロセスを予測シェアが実際のシェアに十分近づくまで繰り返す.