補足資料：離散選択モデルの識別

最終更新：2021 年 5 月 24 日

経済セミナー連載「実証ビジネス・エコノミクス」
第 2 回「需要を制する者はプライシングを制す」（2021 年 5・6 月号）の付録

1 はじめに

この補足資料では、離散選択モデルにおけるパラメタの識別について説明を行う。本連載「実証ビジネス・エコノミクス」を読んでいく中で出てくるであろう、以下のような疑問点に答えることを目標とする。

1. なぜ、アウトサイドグッズからの効用をゼロに基準化するのか？
2. なぜ、個々人の選好ショック ϵ_{ijt} の標準偏差パラメタについて推定しないのか？

なお、本資料の説明は、Train (2009) *Discrete Choice Methods with Simulation* の Chapter 2 に基づいている。詳細についてはそちらを参照されたい[*1]。

2 モデル

連載第 2 回では以下のような効用の定式化に基づく離散選択モデルを考えた。

$$U_{ijt} = \begin{cases} V_{jt} + \epsilon_{ijt} & j = 1, \dots, J \\ V_{0t} + \epsilon_{i0t} & j = 0 \end{cases}.$$

[*1] 同書は著者のウェブサイト（`https://eml.berkeley.edu/books/choice2.html`）で公開されている。

ここで連載とは異なるノーテーションとして V_{jt} を導入しており、これは以下のように定義される[*2]。

$$V_{jt} = \beta^0 + \sum_{k=1}^{K} \beta^k x_{jt}^k - \alpha p_{jt} + \xi_{jt} \; j = 1, \ldots, J \tag{2.1}$$

$$V_{0t} = 0 \tag{2.2}$$

ここで、選択肢 $j = 0$ はアウトサイドグッズと呼び、検討している製品集合のいずれをも購入しないという選択を意味する。

さて、連載ではこのアウトサイドグッズからの効用を $V_{0t} = 0$ という形で基準化している。また、推定においては ϵ_{ijt} の分布に関するパラメタ、特に標準偏差の推定を行っていない。これらの点は、離散選択モデルの性質によるものであり、以下で説明していく。

3 パラメタの識別

離散選択モデルにおいては、各消費者 i は最も高い効用 U_{ijt} が得られるような選択肢 j を選ぶ問題を解く。

$$\max_{j \in \{0,1,\ldots,J\}} U_{ijt}$$

この問題において、消費者は各選択肢から得られる効用 U_{ijt} を比較している。その結果から、個々人の選好ショックの確率的な分布を踏まえると以下のような選択確率が得られる。

$$P_{jt} = \mathrm{Prob}\left(U_{ijt} > U_{ilt}, \forall l \neq j\right) \tag{3.1}$$

$$= \mathrm{Prob}\left(U_{ijt} - U_{ilt} > 0, \forall l \neq j\right) \tag{3.2}$$

この選択確率は選択肢間の効用の差 $U_{ijt} - U_{ilt}$ にしか依存しないことがわかる。この点から、モデルパラメタの推定における 2 つの重要なポイントが得られる。

1. 選択肢間の効用の差を捉えるパラメタしか推定することができないため、特定の選択肢の効用について任意の値に基準化する必要がある。
2. 効用のスケール (水準) を推定することができず、任意のものに設定する必要がある。

前者は、すべての選択肢の効用 U_{ijt} に定数項を足したとしても効用の差は変化しないことに起因する。したがって、ある選択肢の効用を基準として他の選択肢の効用を測る必要がある。この基準化をしばし

[*2] 実はこの V_{jt} は連載で定義した平均効用 δ_{jt} と同じである。ただし、以下の議論は個々人の選好ショック ϵ_{ijt} 以外の項が消費者によって異なっている場合、つまり V_{ijt} として消費者のインデックス i が入ってくる場合でも成立する。例として『経済セミナー』2021 年 8・9 月号に掲載する第 3 回で検討するランダム係数モデルが挙げられる。

ば、ロケーションに関する基準化 (location normalization) と呼ぶ[*3]。これはアウトサイドグッズの基準化と関連している。後者は、効用 U_{ijt} を定数倍しても、最も高い効用が得られる選択肢は変化しないことに起因する。ここからスケールに関する基準化 (scale normalization) が必要となり、これは標準偏差パラメタの推定と関連してくる。これらの点について以下でより詳しく見ていこう。

3.1 効用の差に関する基準化

(3.2) 式を見ると、選択確率 P_{jt} は効用の差 $U_{ijt} - U_{ilt}$ にしか依存しないことがわかる。例えば、新しい効用の定式化として $U_{ijt} + k$ (ただし k は何らかの定数) を考えたとしよう。このとき効用の差をとると k は消えてしまうため、もとの定式化 U_{ijt} から予測される選択確率と全く同じものが得られる。このようなとき、パラメタ k は識別されない (すなわちデータから推定できない) と言う。

この点がアウトサイドグッズからの効用の基準化にどのように影響しているかを考えよう。今、効用の定式化として以下のものを考える。

$$\begin{aligned}\tilde{V}_{jt} &= \beta^0 + \sum_{k=1}^{K} \beta^k x_{jt}^k - \alpha p_{jt} + \xi_{jt} \quad j = 1, \dots, J \\ \tilde{V}_{0t} &= \beta^{outside}\end{aligned}$$

ここで新しいパラメタ $\beta^{outside}$ は、アウトサイドグッズを選択した場合の平均的な効用を示している。

このモデルは最初の定式化 (2.1) と (2.2) よりも一般的と考えられ、追加的なパラメタ $\beta^{outside}$ を推定したいと考えるかもしれない。しかしながら、新たなパラメタ $\beta^{outside}$ はすでに導入されている定数項パラメタ β^0 と別々に推定することは不可能である。というのも、アウトサイドグッズに対する効用差は $\tilde{V}_{jt} - \tilde{V}_{0t} = (\beta^0 - \beta^{outside}) + \sum_{k=1}^{K} \beta^k x_{jt}^k - \alpha p_{jt} + \xi_{jt}$ と書くことができるため、インサイドグッズの定数項 β^0 が大きくなるということは、$\beta^{outside}$ が同じ分だけ小さくなるということちょうど打ち消しあい、モデル上全く同じ予測が得られる。

したがって、モデルのパラメタを推定する際には、ある選択肢からの効用水準を基準化する必要があり、そのためにアウトサイドグッズの効用の基準化を行うのである。上の定式化のもとでこの基準化を

[*3] ある選択肢の効用のレベルを固定するということは、各選択肢からの効用を数直線上にプロットしたときにその選択肢の「位置 (ロケーション)」を固定していることに等しい。これがロケーションに関する基準化と呼ばれる所以である。

行うには、全ての選択肢の効用から $\beta^{outside}$ を引けば良い。すると、以下の定式化が得られる。

$$\begin{aligned}\tilde{V}_{jt} - \beta^{outside} &= (\beta^0 - \beta^{outside}) + \sum_{k=1}^{K} \beta^k x_{jt}^k - \alpha p_{jt} + \xi_{jt} \quad j = 1, \ldots, J \\ \tilde{V}_{0t} - \beta^{outside} &= 0\end{aligned}$$

この定式化は β^0 が $\beta^0 - \beta^{outside}$ に置き換わったことを除いて (2.1) と (2.2) と全く同じであることがわかろう。

以上を踏まえると、アウトサイドグッズからの効用をゼロに基準化したとしてもモデルの一般性が失われないことがわかる。この場合のインサイド・グッズ (すなわち $j = 1, \ldots, J$ の選択肢) における定数項 β^0 は、アウトサイドグッズを選択することの平均的な効用からの相対的な効用値として解釈される。

3.2 効用のスケールに関する基準化

以上では効用に定数を定数項を足し引きしても消費者の選択が変化しない（すなわち選択確率も変化しない）点を見てきた。同様にして、効用を定数倍したとしても消費者の選択肢は変化しないことが言える。というのも、最も高い効用が得られる選択肢は、効用の水準自体には関係がないためである。新たな定式化として、$\tilde{U}_{ijt} = \lambda U_{ijt}$ (ただし、$\lambda > 0$) を考えると、

$$\begin{aligned}\text{Prob}\left(\tilde{U}_{ijt} > \tilde{U}_{ilt}, \forall l \neq j\right) &= \text{Prob}\left(\lambda U_{ijt} > \lambda U_{ilt}, \forall l \neq j\right) \\ &= \text{Prob}\left(U_{ijt} > U_{ilt}, \forall l \neq j\right)\end{aligned}$$

となり、元の定式化と全く同じ予測となることがわかる。

この点を踏まえると、離散選択モデルにおいては効用のスケールを基準化する必要がある。この役割を担っているのが、個々人の選好ショック ϵ_{ijt} の標準偏差に関する基準化である。この点についてロジットモデルを用いて説明する。まず、元の定式化を再掲しよう。

$$U_{ijt} = V_{jt} + \epsilon_{ijt} \quad j = 1, \ldots, J \tag{3.3}$$

$$U_{i0t} = V_{0t} + \epsilon_{i0t} \tag{3.4}$$

ここで ϵ_{ijt} は i.i.d. の第一種極値分布に従い、また、その分散を $Var(\epsilon_{ijt}) = \frac{\pi^2}{6}\sigma^2$ とする。なお $\pi^2/6$ が出てくる理由については後述する。

このモデルは、効用を σ で割った定式化と全く同じ予測となる。

$$\tilde{U}_{ijt} = \frac{U_{ijt}}{\sigma} = \frac{V_{jt}}{\sigma} + \tilde{\epsilon}_{ijt} \quad j = 1, \ldots, J$$
$$\tilde{U}_{i0t} = \frac{U_{i0t}}{\sigma} = \frac{V_{0t}}{\sigma} + \tilde{\epsilon}_{i0t}$$

ここで、$\tilde{\epsilon}_{ijt} = \epsilon_{ijt}/\sigma$ であり、したがって $Var(\tilde{\epsilon}_{ijt}) = \frac{\pi^2}{6}$ となる。

連載第 1 回及び第 2 回では明示的に書いていなかった点であるが、ロジット確率が導出される場合の第一種極値分布のショックの分散は $\frac{\pi^2}{6}$ となっている。ここから、ロジットモデルの選択確率は

$$\text{Prob}(choose\ j) = \frac{\exp(\frac{V_{jt}}{\sigma})}{\exp(\frac{V_{0t}}{\sigma}) + \sum_{k=1}^{J} \exp(\frac{V_{kt}}{\sigma})}$$

となる。この点をみてわかるように、σ と V_{jt} は比率の形で入ってくるため、それぞれの値を別個に推定することは不可能である。もし $V_{jt} = \beta' X_{jt}$ と書くと、推定可能なパラメタは β/σ という比率になる。

以上を踏まえると、V_{jt} の中に現れる係数と標準偏差パラメタ σ を別個に推定することは不可能となる。結果、効用のスケールの固定として $\sigma = 1$ を置く、すなわち選好ショック ϵ_{ijt} の分散を $\frac{\pi^2}{6}$ に基準化する必要が出てくるのである。これが、連載第 1 回及び第 2 回のロジットモデルの導出において暗示的に用いていた議論となる。