

多数の選択肢を有する選択行動への非集計行動モデルの適用可能性の検討

—観光目的地選択行動をケーススタディーとして—

非会員 東京大学大学院 *福田 大輔 FUKUDA Daisuke
 01601460 東京大学大学院 森地 茂 MORICHI Shigeru

1. はじめに

近年、パラメータ推定における数値積分法およびシミュレーション法の発展と共に、誤差項の構造を精緻化した非集計行動モデルの開発が精力的に行われている。こうしたモデルが実用に至るためには、様々な選択行動に対する適用事例の蓄積が必要である。このような認識に基づき、本研究では多数の選択肢を有する行動に対する精緻化された非集計行動モデルの適用可能性を検討することを目的とする。ここでは特に選択肢の誤差構造に着目し、誤差項の分散不均一を表現する Heteroscedastic-Extreme-Value モデル、及び選択肢間の類似性を表現するモデルとして Mixed-Logit モデルの2つを取り上げる。それらのモデルの特性を、観光目的地選択行動のデータを用いた実証分析を通じて把握することを目指す。

2. 選択肢の誤差構造の精緻化

(1) Heteroscedastic-Extreme-Value(HEV)モデル

HEV モデル¹⁾は、誤差項間の独立の仮定は保持されているものの、分散の大きさが各選択肢によって一般には異なることを想定した離散選択モデルである。これにより、HEV モデルは I.I.A. 性質を厳密には保持せず、各選択肢に対する認識のあいまいさの度合いを分散の大きさの相違として表現することが可能である。選択確率は次のように定式化される。

$$P_{n,i} = \int_{-\infty}^{\infty} \prod_{k=i,k \in C_n} F[\theta_k (V_{n,i} - V_{n,k} + \epsilon_i)] \theta_i f(\theta_i \epsilon_i) d\epsilon_i \dots (1)$$

ここで、 $F(\theta_i \epsilon_i) = \exp(-\exp(-\theta_i \epsilon_i))$: 誤差項 ϵ_i の累積密度関数、 $f(\theta_i \epsilon_i)$: 誤差項 ϵ_i の確率密度関数、 θ_i : 選択肢 i の誤差項のスケールパラメータ、 $V_{n,i} = (\beta' \cdot X_{n,i})$: 確定効用項 (変数は表1を参照) である。また、誤差項の分散は $\pi^2/6\theta_i^2$ で与えられるため、 θ_i の値が小さい選択肢ほど、誤差のばらつきが

大きい選択肢である。ここで、あえて適用事例に対する行動論的な解釈を行えば、認識が曖昧である目的地ほど、 θ_i の値が小さいと解釈することもできる。

(2) Mixed-Logit モデル

Mixed-Logit モデル²⁾は、線形効用関数を持つロジットモデルにおいて、その説明変数のパラメータがランダムに変動することを仮定したモデルであり、説明変数の設定を特定化することで、選択肢間の類似性を明示的に表現することが可能である。本研究では、Brownstone and Train(1998)³⁾ を参考に、効用関数を次式のように設定する。

$$U_{n,i} = \beta' \cdot X_{n,i} + \mu' \cdot Z_i + \epsilon_i \dots (2)$$

ここで、 μ : 平均 0 の同時正規分布にしたがう要素数 $(J_n - 1) \times J_n / 2$ 個のランダムパラメータベクトル (U_n : 個人 n の選択肢集合に含まれる選択肢数)、 Z_i : 選択肢 i と他の選択肢間の類似度ベクトル、 ϵ_i : 誤差項 (i.i.d. Gumbel) である。第2項 $\mu' \cdot Z_i$ の存在により、選択肢間の類似性が明示的に表現可能となるが、ここでは類似度ベクトル Z_i を式(3)のように特定化し、さらに、 μ の各要素が独立で同一の分散 ω^2 を有すると仮定すれば、選択肢 i と選択肢 j の共分散は式(4)のように簡潔に表される。

$$Z_i = (z_{i,2}, z_{i,3}, \dots, z_{i,J_n}, z_{2,3}, \dots, z_{2,J_n}, \dots, z_{J_n-1,J_n})' \dots (3)$$

$$z_{p,q} = \begin{cases} w_{p,q} & \text{if } p = i \text{ or } q = i \\ 0 & \text{otherwise} \end{cases}$$

$$E[(\mu' \cdot Z_i + \epsilon_{n,i}) (\mu' \cdot Z_j + \epsilon_{n,j})] = w_{ij}^2 \cdot \omega^2 \dots (4)$$

ここで、 w_{ij} : i 番目の目的地と j 番目の目的地の類似度指標であり、目的地間の距離の逆数をもってその値とした。なお選択確率は、次式で表される。

$$P_{n,i} = \int_{-\infty}^{\infty} \frac{\exp(\beta' \cdot X_{n,i} + \mu \cdot Z_i)}{\sum_{j \in C_n} \exp(\beta' \cdot X_{n,j} + \mu \cdot Z_j)} g(\mu; \omega) d\mu \dots (5)$$

3. 実証分析

用いるデータは、全国観光交通実態調査(1992)の自動車日帰り旅行の目的地選択データである。全国を9つの出発地ブロックに分類し、ブロックごとにモデル化を行う。なお基本選択肢集合の設定方法、使用する説明変数は、Okamoto *et al.*(1995)⁴⁾ に依拠する。パラメータの推定は、式(1)、(5)共に積分が閉じていないため、HEV モデルは、Gauss-Laguerre 積分法を用いて式(1)を近似してから、また Mixed-Logit モデルはシミュレーション法を用いて最尤推定を行う。

紙面の都合上、中国エリア居住者 (1,158 サンプル) に対する説明変数パラメータ $\beta_1 \sim \beta_6$ の推定結果のみを表 1 に示す。ほとんどの推定値は統計的に十分有意であるが、値の大きさはモデル間でかなり異なっている。特に、所要時間や交通費用といった、プロジェクト評価上重要なパラメータも値が異なっており、誤差項に関する十分な配慮の必要性が再確認される。また、表 2 の AIC 等の比較結果からは、HEV 及び Mixed-Logit を採用することにより推計精度が大きく向上していることが確認できる。

ところで、Mixed-Logit モデルのパラメータ推定では先述の ω の値も推定され、中国エリアの場合この値は 10.8×10^{-4} (t 値 : 1.12) であり、統計的にはそれほど有意ではなかった。しかし、選択肢間の類似性を考慮することによりモデルの説明力は大きく向上しており、多数の代替案を有する選択行動に対しても、何らかの類似度指標を用いて誤差項の構造化を行うことにより、モデルの精度はまだ向上する可能性が示唆される。

表 1: パラメータ推定結果 (カッコ内 t 値)

パラメータ	Logit	HEV	Mixed-Logit
所要時間(分)	-9.54×10^{-3} (-20.12)	-11.41×10^{-3} (-14.47)	-9.61×10^{-3} (-17.92)
交通費用(円) / ln(年収[万円])	-5.55×10^{-4} (-3.81)	-7.53×10^{-4} (-3.25)	-6.89×10^{-4} (-3.99)
見物・鑑賞	7.76×10^{-1} (10.42)	9.65×10^{-1} (4.93)	3.64×10^{-1} (4.29)
海レジャー	14.69×10^{-1} (7.09)	12.94×10^{-1} (6.52)	12.93×10^{-1} (7.26)
温泉	5.06×10^{-1} (2.65)	8.26×10^{-1} (3.08)	3.98×10^{-1} (1.71)
山レジャー	8.59×10^{-1} (6.85)	11.7×10^{-1} (5.75)	7.69×10^{-1} (5.93)

表 2: 各種統計量の比較

各種統計量	Logit	HEV	Mixed-Logit
最大対数尤度	-2296.6	-2173.3	-2060.5
自由度調整済尤度比	0.227	0.268	0.306
χ^2 統計量	1347.2	1593.8	1819.5
AIC	4605.2	4382.6	4135.0
目的地数	13		
初期対数尤度	-2970.2		

次に、誤差項の分散不均一性を検討するために HEV モデルを関東エリア居住者 (2,464 サンプル、基本選択肢 20 個) のデータに対して適用した。紙面の都合上、ここでは考察結果のみを記すが、推定された各目的地の標準偏差 ($\pi/\sqrt{6} \theta_i$) の大きさにはかなりのばらつきが見られ、通常のロジットモデルの分散均一という仮定がいかに強力なものであるかが確認された。また、標準偏差の値と東京から各目的地までの距離などとの対応に関しても考察を試みたが、明確な関係 (例えば、発地からの距離が大きくなるにつれて標準偏差の値も大きくなるなど) を見出すことはできなかった。今後の課題としたい。

4. おわりに

今回は分散不均一と選択肢間の類似性の問題を別個に考えたが、Mixed-Logit モデルをさらに精緻化すれば、これらを同時に取り扱うことが可能である。今後はこうした統合化の方向性を探ると同時に、選択肢集合の設定方法や、1 つの選択肢として設定されるエリアの大きさの妥当性などといった、選択問題に対するより根本的な問題の検討が必要である。

参考文献

- 1) Bhat, C. R. (1995) A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B*, Vol.29, pp.471-483.
- 2) McFadden, D. and Train, K. (1997) Mixed multinomial logit models for discrete response. *Working Paper*, Department of Economics, University of California at Berkeley, CA, USA.
- 3) Brownstone, D. and Train, K. (1998) Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, Vol.89, 1-2, pp.109-129.
- 4) Okamoto, N., Yai, T., Morichi, S. and Nishimura, T. (1995) A study on regional difference of recreational destination choice behavior, *Journal of the Eastern Asia Society for Transportation Studies*, Vol.1, No.1, pp.351-360.