

評価属性の差異を考慮した離散型選択モデルに関する研究*

Discrete Choice Model with Latent Classes Considering Heterogeneity of Attribute Sets*

倉内慎也**, 森川高行***, 祖父江誠二****

By Shinya KURAUCHI**, Takayuki MORIKAWA***, and Seiji SOBUE****

1. はじめに

ロジットモデルに代表される多属性効用理論に基づく離散型選択モデルでは、意思決定者が想起する選択肢集合、及び選択肢を評価する際に着目する属性の集合を、分析者が確定的に決定する必要がある。しかし、多くの場合、個々の意思決定者の持つ真の選択肢集合及び評価に際して着目する属性の集合は、分析者にとって不明である。これまで、選択肢集合に着目した研究は数多くなされており、誤った選択肢集合を特定した場合にはパラメータ推定値に重大なバイアスをもたらすことが実証され¹⁾、選択肢集合の不確実性を明示的に考慮したモデルが開発されている^{2), 3)}。一方、意思決定者が選択肢の評価に際して着目する属性については、主に購買行動を対象とした消費者行動研究の分野において盛んに議論されており、そこでは、人間の情報処理能力に限界があるため、ごく限られた重要な属性のみを評価して選択を行なうという非補償型意思決定方略を採用する傾向が強いことが数多く報告されている⁴⁾。

そのような主張に反して、交通行動分析においては、属性の線形形で表される効用関数が暗黙裡に用いられており、これは主に以下の2点に関して問題があると考えられる。第一に、線型効用関数を用いることは、「意思決定者はそこに含まれる全ての代替案属性を評価している」と仮定することを意味しており、目的地選択に代表されるように対象問題の選択肢数や選択肢を構成する属性の数が多い場合には、行動論的に見て不適切な場合があると考えられる。第二に、線型効用関数を用いて需要予測を行なった場合、対象サンプルに含まれる全個人の選択確率はサービスレベルの向上に対して必ず増加するが、

「約束の時間に間に合うのであれば費用は問わない」といったような個人については、選択確率は不変であり、そのようなサンプルが少なからず存在する場合には、当該施策の効果を過大に評価する危険性がある。

以上の問題を踏まえ、本研究では、意思決定者が選択肢の評価に際して着目する属性集合の不確実性を考慮した離散型選択モデルを構築し、事例研究を通じて、その推定特性やモデルの挙動に関して検討することを目的とする。

2. PAS (Probabilistic Attribute Set) モデル

本研究では、評価属性集合の不確実性を明示的に考慮するために、個人間で異なる選択肢集合を確率的に記述した Manski の確率的選択肢集合 (PCS: Probabilistic Choice Set) モデル²⁾の考え方にに基づき、選択肢集合をそのまま属性集合に置き換えることによってモデル化を行なった。すなわち、個人 n が選択肢 i をする確率 $P_n(i)$ は下式のように表すことができる。

$$P_n(i) = \sum_{A \in G} P_n(i|A) Q_n(A) \quad (1)$$

ここに、 $P_n(i|A)$: 個人 n の評価属性集合が A であったときの代替案 i の選択確率、 $Q_n(A)$: 個人 n の評価属性集合が A である確率、 G : 属性のすべての部分集合の集合、である。

このモデルは、 $Q_n(A)$ で表される評価属性集合の形成と、 $P_n(i|A)$ で表される、評価属性集合が所与の元での選択行動、という2段階の構成になっており、起り得る全ての属性集合を考慮する、すなわち属性集合の周辺確率を評価することによって選択確率を導出している。ここで、PCS モデルにおいては、選択肢集合が空集合になることを避けるために、空集合が生起する確率を用いて、全選択肢の選択確率の

* キーワード: 交通行動分析, 経路選択
** 正会員 工修 名古屋大学大学院工学研究科地圏環境工学専攻
*** 正会員 Ph.D. 名古屋大学大学院環境学研究科都市環境学専攻
**** 学生員 名古屋大学大学院環境学研究科都市環境学専攻
(〒464-8603 名古屋市中区不老町, tel:052-789-3565, fax:052-789-3738)

合計が 1 となるような補正を必要とするが、確率的属性集合モデル（以下、PCS モデルにちなんで PAS: Probabilistic Attribute Set モデルと呼称）では、属性集合が空集合になるような場合も、「意思決定者は選択肢の属性を全く考慮せずにランダムに決定を行う」という可能性を否定できないと考え、本研究では起こり得る全ての属性集合を評価するような構造を採用した。

ここで、個人 n の評価属性集合が A であったときの代替案 i の効用関数 $U_n(i|A)$ は、属性 k が個人 n の評価属性集合 A に含まれる場合は β_k 、それ以外は 0 となるような特殊な未知パラメータ β_k を用いることにより、従来の線型効用関数と全く等価な形で記述することができる。

$$U_n(i|A) = V_n(i|A) + \varepsilon_n(i|A) = \sum_k \beta_k X_{kin} + \varepsilon_n(i|A) \quad (2)$$

$$\beta_k = \begin{cases} \beta_k : & k \in A \\ 0 : & \text{otherwise} \end{cases} \quad (3)$$

ここに、 $V_n(i|A)$ ：個人 n の評価属性集合が A であったときの代替案 i の効用関数の確定項、 $\varepsilon_n(i|A)$ ：個人 n の評価属性集合が A であったときの代替案 i の効用関数の誤差項、 X_{kin} ：個人 n が直面している代替案 i に関する属性 k の値、である。式 (2) において、誤差項 $\varepsilon_n(i|A)$ に i.i.d. ガンベル分布を仮定すれば、式 (1) の $P_n(i|A)$ は、通常のプロジットモデルの形で表される。

$$P_n(i|A) = \frac{\exp(V_n(i|A))}{\sum_j \exp(V_n(j|A))} \quad (4)$$

個人 n の評価属性集合が A である確率 $Q_n(A)$ については、PCS モデルにおける Swait らの研究³⁾と同様に、評価属性集合は各属性を評価するか否かの二項選択確率の積で表されるものとして定式化する。

$$Q_n(A) = \prod_{k \in A} q_n(k) \prod_{k \notin A} (1 - q_n(k)) \quad (5)$$

$$q_n(k) = \frac{\exp(H_{kn})}{1 + \exp(H_{kn})} \quad (6)$$

ここに、 $q_n(k)$ ：個人 n が属性 k を考慮して評価を行なう確率、 H_{kn} ：個人 n が属性 k を考慮して評価を行なう確率を規定するメンバーシップ関数、である。

最終的に、式 (2) ～ (6) を式 (1) に代入して得られる選択確率を各個人の尤度として最尤推定法を行なうことにより、効用関数及びメンバーシップ関数に含まれる未知パラメータを推定することができる。

3. 事例研究

(1) データの概略

本研究では、建設省（現国土交通省）が実施した平成11年全国都市パーソントリップ調査の付帯調査データのうち、買い物・レジャー交通に関して、一般道と有料道の2経路のみが存在するような仮想状況下での経路選択を尋ねたSPデータを用いた。表1に示す設定条件の下でいずれの経路を選択するかを尋ね、1個人当たり12個の選択データが得られている。

本データは、代替案数と属性数が共に少数であるため、1. で述べたような問題意識と必ずしも整合していない部分がある。しかし、1) SPデータであるため、属性間のトレードオフが明瞭であると同時に、所要時間等のLOS変数に内在する不確実性の問題が比較的少ない、2) データに含まれる被験者数（15379人）が多く、かつ複数地域（73都市）の居住者を対象としているため、意思決定の多様性が大きいと考えられる、等の理由から、モデル特性の考察に主眼を置く本研究に適していると判断し、事例研究に用いるデータとして採用した。

表－1 SP調査におけるサービスレベル変数

一般道		有料道	
所要時間	料金	所要時間	料金
40 分	0 円	30 分	100, 300, 500, 700 円
50 分	0 円	30 分	200, 400, 600, 800 円
60 分	0 円	30 分	300, 600, 900, 1200 円

(2) モデルの推定結果と考察

PAS モデルの特性を既存モデルと比較検討するために、以下の4つのモデルを別途推定した。

- ・ MNL-1：定数項と LOS 変数のみを説明変数とした多項ロジットモデル
- ・ MNL-2：嗜好の異質性を表すために MNL-1 に観測可能な個人属性を付加したモデル
- ・ RC：ランダム係数モデル⁵⁾（定数項と LOS 変数のパラメータに独立な標準正規分布を仮定）

- ・ L-LC: 2つの異なる線形効用関数を仮定した潜在クラスモデル⁶⁾

推定結果を表-2に示す。まず、AIC の値から、モデルの適合度は、RC モデルが卓越して優れており、次いで PAS モデル、L-LC モデル、MNL モデルの順になっている。これは、本データでは、1) 嗜好の分布が単峰型に近く、かつ、2) 観測不可能な意思決定者の特性による嗜好の異質性が卓越しており、その点において、3) 1 個人当たり 12 個の選択結果の同時出現確率を尤度として定式化した本研究では、唯一 RC モデルが同一個人に共通な嗜好を反映できるモデル構造を持っている、ためであると考えられる。しかしながら、RC モデルは、多重積分を必要とするため、他のモデルと比較して計算負荷が大きく、属性数が増えた場合を考慮すると、実用性に欠けると思われる。一方、PAS モデルの推定時間は実用範囲内にあると判断でき、また、L-LC モデルにおいて各潜在クラスがどのような嗜好を持つクラスであるのかを解釈するのが一般的に困難であるのに対して、PAS モデルでは、例えば、「高齢者は所要時間をあまり評価しない傾向にある」など、その意味付けが明確であるという利点がある。加えて、表-3に示すように、「サンプルの約 65%が全ての属性を評価する傾向がある」など、各個人が意思決定に際して考慮する属性集合を評価できるため、政策決定に際して多くの有用な情報を提供できるものと期待される。

(3) モデルの挙動分析

前節で推定したモデルを用いて、政策変数の変化に対する集計シェアの挙動分析を行なった(図-1, 2)。なお、集計化にあたっては、各モデルに含まれる誤差項の分布に従い、1 個人当たり 10000 個の乱数を発生させ、その平均値を当該個人の選択確率とみなして集計シェアを計算した。また、各モデルの挙動が観測可能な個人属性によってどのように変化するかを分析するために、会社勤務の男性(セグメント 1: 家族 4 人, 50 歳, 管理職, 専用自動車あり)と独居老人(セグメント 2: 男性, 70 歳, 無職, 専用自動車なし)に着目して、LOS 変数の変化に対する選択確率の推移を計算した(図-3, 4)。

いずれの場合においても、PAS 以外のモデルではシェア(選択確率)の変化が比較的緩やかであるの

に対して、PAS モデルでは、LOS 変数の変化率が少ない場合にはシェア(選択確率)はあまり変わらず、その後急激に変化した後、再びシェア(選択確率)の変化が緩やかになっている。これは、PAS モデルは、線形効用関数のパラメータが 0 になる、すなわちある属性が全く評価されないという非補償型意思決定方略を含んでおり、かつその意思決定方略への帰属確率が 0 でないためであり、実際の意思決定状況において非補償型の意思決定方略が頻繁に観測されるという多くの研究報告⁴⁾とも整合している。

4. おわりに

本研究では、意思決定者は必ずしも線形効用関数に含まれる全ての代替案属性を考慮しているとは限らないとの認識の下、Manski の提唱した二段階モデル²⁾の枠組みに基づいて、個人間で異なる評価属性集合の差異を考慮した離散型選択モデルを構築した。また、事例研究において、既存モデルとの比較を行ない、提案モデルは、解釈が明確である、政策決定に際して多くの有用な情報を提供し得る、計算負荷は実用範囲内にある、モデルの挙動が実社会での直感や既存研究と整合的である、等の利点を明らかにした。しかし、本研究はいまだ途上段階であり多くの問題点を抱えている。引き続き事例研究を重ね、モデルの特性を精査すると共に、どのような選択問題に有効であるのか等、更なる検討が不可欠である。

最後に、分析にあたり、全国都市パーソントリップ調査ワーキング(座長: 東京大学・原田昇教授, 事務局: 国土交通省・国土技術政策総合研究所)において貴重な助言を頂戴した。ここに記して、深謝の意を表します。

参考文献

- 1) Williams, H. C. W. L. and Ortuzar, J. D.: Behavioural theories of dispersion and the mis-specification of travel demand models, *Transportation Research*, 16B, pp.91-102, 1982.
- 2) Manski, C.: The structure of random utility models, *Theory and Decision* 8, pp.229-254, 1977.
- 3) Swait, J. and Ben-Akiva, M.: Incorporating random constraints in discrete models of choice set generation, *Transportation Research*, 21B, pp.91-102, 1987.
- 4) Payne, J., Bettman, J. and Johnson, E.: *The Adaptive Decision Maker*, Cambridge University Press, New York, 1993.
- 5) Daganzo, C.: *Multinomial Probit*, Academic Press, 1979.
- 6) Sasaki, K., T. Morikawa, S. Kawakami: A Discrete Choice Model with Taste Heterogeneity Using SP, RP and Attribute Importance Ratings, *Proc. of 8th World Conf. on Transport Research*, 1998.
- 7) 山内二郎, 森口繁一, 一松信編: 電子計算機のための数値計算法 I, 培風館, 1965.

表-2 各モデルの推定結果 (括弧内はt値)

説明変数	MNL-1	MNL-2	RC 注1	L-LC	
				《Class 1》	《Class 2》
一般道定数項	2.94 (149.4)	3.04 (136.0)	3.83 (155.5)	3.23 (66.5)	1.95 (15.3)
料金 (千円)	-3.47 (-143.0)	-3.51 (-143.5)	-5.21 (-163.5)	-4.11 (-91.0)	-7.66 (-10.9)
所要時間 (時間)	-10.5 (-173.9)	-10.6 (-174.2)	-14.4 (-192.9)	-13.7 (-69.4)	-7.06 (-15.3)
				《Class 1 のメンバーシップ関数》 注3	
定数項				0.726 (9.3)	
世帯人数 2人以上がミ		-0.0317 (-1.8) 注2		0.133 (2.9)	
男性がミ		-0.000332 (-0.02) 注2		-0.162 (-3.7)	
30歳以下がミ		0.108 (6.3) 注2		-0.218 (-5.0)	
60歳以上がミ		0.341 (21.6) 注2		-0.774 (-20.5)	
管理職がミ		-0.415 (-15.1) 注2		1.261 (8.7)	
専用自動車保有がミ		-0.266 (-19.9) 注2		0.660 (17.6)	
サンプル数	15379	15379	15379	15379	
パラメータ数	3	9	3	13	
AIC	166573	165025	61580	164348	
計算時間 (分) 注4	0.2	3.2	616.27	13.7	
説明変数	PAS				
一般道定数項	4.04 (20.7)				
料金 (千円)	-6.34 (-25.5)				
所要時間 (時間)	-17.8 (-22.7)				
	《料金のメンバーシップ関数》		《所要時間のメンバーシップ関数》	《その他要因のメンバーシップ関数》	
定数項	2.739 (15.9)	1.170 (20.2)		3.241 (7.6)	
世帯人数 2人以上がミ	0.164 (1.1)	0.237 (5.0)		0.135 (1.5)	
男性がミ	0.499 (2.4)	-0.376 (-8.0)		-1.106 (-4.1)	
30歳以下がミ	1.048 (2.6)	-0.041 (-0.6)		-0.019 (-0.1)	
60歳以上がミ	-0.736 (-5.7)	-0.830 (-20.4)		0.253 (1.5)	
管理職がミ	-0.915 (-4.3)	0.616 (6.3)		-0.221 (-1.6)	
専用自動車保有がミ	-0.176 (-1.7)	0.524 (14.0)		-0.171 (-2.0)	
サンプル数	15379				
パラメータ数	24				
AIC	164015				
計算時間 (分) 注4	116.14				

注1) 積分計算については Gauss-Legendre の公式⁷⁾に基づいて数値的に評価、注2) 一般道の効用関数に含まれる説明変数

注3) ロジット型メンバーシップ関数を用い Class 2 のメンバーシップ関数の確定項をゼロと固定、注4) 統計計算用言語 GAUSS 3.2 を用いて CPU: Pentium IV 1.4GHz の市販の PC により推定

表-3 PAS モデルから計算される評価属性集合の平均生成確率

評価属性集合	{なし}	{その他要因}	{料金}	{所要時間}
生成確率	0.1	1.5	1.6	0.3
評価属性集合	{その他要因, 料金}	{その他要因, 所要時間}	{料金, 所要時間}	{その他要因, 料金, 所要時間}
生成確率	4.1	21.6	5.1	65.7

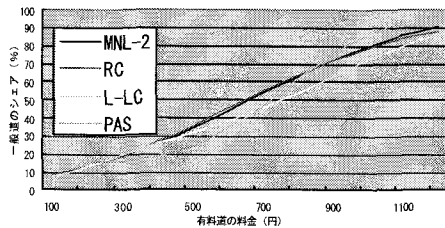


図-1 各モデルの集計予測シェア (料金)

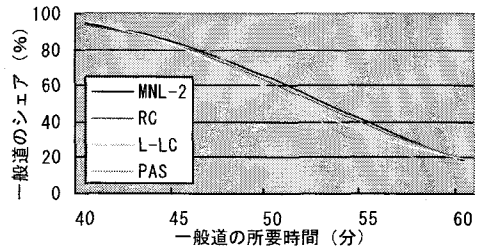


図-2 各モデルの集計予測シェア (所要時間)

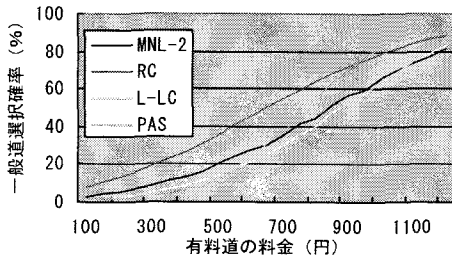


図-3 選択確率の推移 (セグメント1)

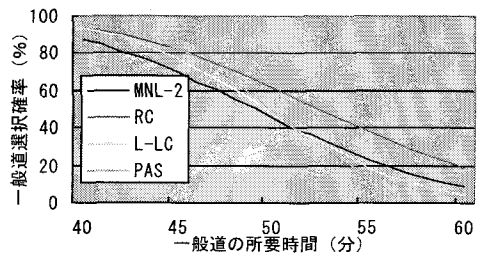


図-4 選択確率の推移 (セグメント2)