

交通行動分析における内的情報を用いたセグメンテーションに関する研究

J R 東海 岡崎真人
名古屋大学 河上省吾
名古屋大学 佐々木邦明

1. はじめに

これまで非集計離散型選択モデルに個人の異質性を導入する簡便な手法として、マーケットセグメンテーションがある。一般にセグメンテーション手法は大きく分けて、外的基準によるもの、内的基準によるものの二つに分けられる。外的基準によるものとは、通勤・通学、買物レジャーなどの目的別や、就業者・非就業者などの観測可能な外的基準によって母集団を分割し、それぞれの集団ごとに異なるパラメータを持つことを仮定しパラメータを推定する方法である。この手法は容易でありかつ、直感的に理解しやすいため一般に用いられることが多い。しかし近年の価値観の多様化が問題となってきた場においては、このアプローチではそこで問題となっている嗜好の多様性と外的基準との直接的な因果関係が把握できないなどの限界があり、個人の嗜好の違いに基づいた内的情報によるセグメンテーション手法の必要性が増大している。本研究は意思決定者の嗜好の違いを仮想の状況における選択結果や意識データを用いて抽出し、それに基づいてセグメンテーションを行う方法について検討したものである。

フレームワークとしては、個人から得られた複数の選択データによって個人ごとにパラメータを推定し、それによってセグメンテーションを行う方法と、好みの違いを直接尋ねた態度データより、潜在的な態度変数を推定し、それを用いてセグメンテーションを行う方法からなる2種類の方法を用い、両者の比較検討を行う。

2. 内的情報を用いたセグメンテーション

1) 個人パラメータを用いた方法

SPデータは仮想の状況に対しての選好の意思表示であるから、RPデータと違い質問の状況を変化させることによって一人の被験者から同時に複数の回答を得ることができる。この特長を生かして得られた複数選択データより各個人ごとにロジットモデルを推定し、そこで得られたパラメータを用いてセグメンテーションを行い、各セグメントごとにロジットモデルを推定する。ただし、この手法は各個人から得られるデータ数は一般に有効サンプル数と負の相関関係にあり、一個人から得る情報を多くすると総サンプル数が減少してしまうため、一人のサンプルから多くのデータを得ることはできず、個人パラメータの有意性に問題が生じる。

2) 潜在的態度変数による方法

アンケートによって各個人から選択において代替案のどの属性を重視するのかを尋ね、そのデータを指標としてLISRELモデルを用いて潜在的な態度変

数を推定し、そこで得られた値を用いてセグメンテーションを行い、各セグメントごとにロジットモデルを推定する。

3. 事例研究

1) 使用データ

本研究で用いるデータは、1992年10月に千葉県館山を訪れた人を対象に、無作為にアンケート調査票を配布して郵送で回収されたものである。アンケートの内容としては、主に次の4種類に分類することができる。

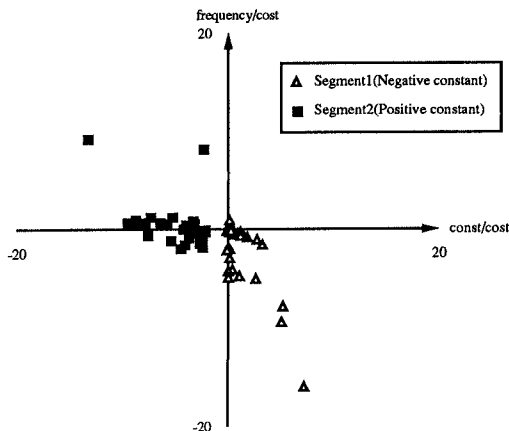
1. 個人の社会経済属性に関するデータ
2. 今回の旅行の手段、経路、旅行内容に関するデータ(RPデータ)
3. 選択時に重視する項目を尋ねた態度データ
4. 新規航路の選択に関するデータ(SPデータ)

神奈川県久里浜一館山間を仮想の新規高速船(フェリー)が運航されているという仮定のもと、新規高速船の所要時間は一定とし、各個人に6パターンの料金と運行頻度の組み合わせを設定し、新規高速船利用の2項選択データを得ている。

全サンプル852人のうち、有効な回答が得られたサンプルは211人であった。個人パラメータ推定に必要なデータが得られたサンプルは84人であった。

2) 個人パラメータの推定結果と考察

各個人ごとにロジットモデルを推定し、得られた定数項、コスト、運行頻度の3つのパラメータを、各個人の誤差項のスケールの違いの影響を除くために、定数項及び頻度のパラメータをコストのパラメータで除した。図-1にその分布状況と平方ユークリッド距離を用いてセグメンテーションを行った結果を示す。図を見ると、ほぼ定数項/コスト軸の正負によって分けられていることがわかる。コストに対するパラメータがほぼ全員が負であったことを考えると、セグメント1は定数項が負であるグループ、セグメント2は定数項が正であるグループであると考えられる。ここで各セグメントごとに2項ロジットモデルを推定した結果を表-1に示す。推定されたパラメータはどれも妥当な符号をもっており、尤度比検定によって、両セグメントのパラメータベクトルの等価性は有意水準5%で却下された。またその全体の χ^2 もセグメントを行うことにより向上している。ここで各グループを構成しているサンプルを詳しく見ていくと、セグメント1はほとんどが鉄道利用者であり、セグメント2は全員が車利用者であることが判明した。もし、嗜好の違いが現在の利用交通手段で説明できるなら、個人パラメータを推定しなくても同様の結果が得られることにな



図－1 個人パラメータの分布図

表－1 各セグメントごとの推定結果

	Pooled	Segment1	Segment2
定数	-1.69 (-6.8)	-3.55 (-6.6)	1.28 (2.0)
費用	-0.312 (-7.2)	-0.340 (-3.9)	-0.704 (-5.4)
頻度	1.91 (8.3)	4.02 (6.1)	1.06 (3.9)
時間	-0.0649(-2.1)	-0.103 (-1.7)	-0.0293(-0.8)
N	492	240	252
ρ^2	0.255	0.325	

るため、利用交通手段別に同様のモデルを推定した。その結果、車利用者の推定されたパラメータはセグメント2と異なりセグメント1と2の中間をとるような値になった。また鉄道利用者グループのパラメータはセグメント1とほぼ同じものとなった。このことから車利用者に関しては嗜好の違いを個人パラメータから抽出できたものと考えられ、鉄道利用者の結果からセグメント1の安定性が確認できた。このことはセグメント1の嗜好の同質性の傍証となると考えられるが、セグメンテーションの基準となった定数項は、費用、頻度で説明されない個人の好みを示しており、本研究で用いたSPデータが各利用手段別に同一のものであり現在利用している交通手段との比較になっているため、鉄道利用者にはフェリーターミナルへの経路情報の不足などにより、定数項が過小評価されたためかも知れない。

3) 潜在的態度変数によるセグメンテーションの推定結果と考察

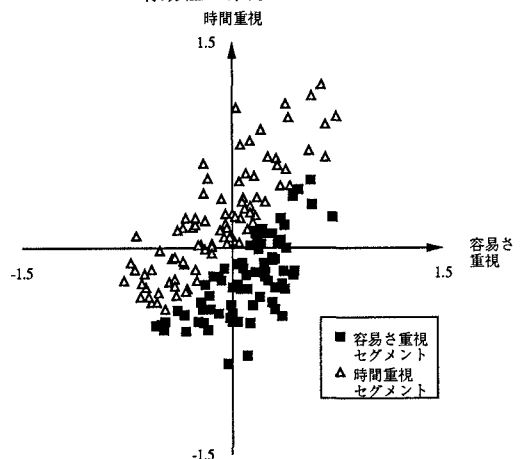
交通手段選択時に重視する項目を尋ねたデータと個人属性からLISRELモデルにより、所用時間重視変数と、利用の容易さの2つの潜在変数を定式化した。その分布状況を図－2に示す。個人の嗜好の違いが効用パラメータに与える影響を考え、セグメントを2つの潜在変数の大小関係によって分割した。つまり図－2の左上が時間重視セグメント、右下が利用の容易さ重視セグメントとなる。ここで各セグ

メントごとにロジットモデルを推定するのであるが、前節の結果より鉄道利用者特有のバイアスを考慮して、鉄道利用ダミーを導入した結果が表－2に示されている。ここでそれぞれのセグメントの時間価値を比較すると、時間重視セグメントは2100円/時、利用の容易さ重視セグメントは1300円/時となり大きな差が見られる。

4. まとめ

本研究の成果を以下の3点にまとめた。

- 1) 個人パラメータに基づくアプローチからは個人の異質性を反映したセグメンテーションを行うことができ、個人パラメータによる方法の有効性を確認できた。
- 2) 複数SPデータによる個人パラメータの推定は、データ数の制約などから得られるパラメータの数に限界があり、推定したモデルの信頼性に問題があり、得られた結果はバイアスを含んでいる可能性が示された。
- 3) 潜在的態度変数によるセグメンテーションから、予想される特質を反映したセグメンテーションが行え、態度変数によるセグメンテーションの有効性が確認できた。



図－2 態度変数の分布状況

表－2 各セグメントごとの推定結果

	Pooled	Segment1	Segment2
定数	-0.684 (-2.3)	-0.469 (-1.2)	-0.781 (-1.7)
費用	-0.426 (-8.7)	-0.380 (-6.7)	-0.512 (-5.7)
頻度	1.39 (9.2)	1.11 (5.9)	1.75 (7.4)
時間	-0.0753(-3.1)	-0.0262(-0.8)	-0.126 (-3.7)
RP	-0.827 (-3.0)	-0.975 (-2.5)	-1.02 (-2.4)
N	819	422	397
ρ^2	0.273	0.275	