

# 潜在クラスモデルを用いた 宿泊観光旅行回数の基礎的分析

古屋 秀樹<sup>1</sup>・全 相鎮<sup>2</sup>

<sup>1</sup>正会員 東洋大学教授 国際地域学部国際観光学科 (〒112-8606 東京都文京区白山5-28-20)

E-mail: furuya@toyo.jp

<sup>2</sup>非会員 日本観光振興協会総合研究所 (〒105-0001 東京都港区虎ノ門3-1-1)

E-mail: s-zen@nihon-kankou.or.jp

本研究は、平成24年度に実施した「国民の観光に関する動向調査」データを用いて、個人の年間宿泊観光旅行回数を潜在クラスモデルを用いて分析する。分析では、旅行回数を推定する関数を同定するとともに、同一の旅行回数関数を有するセグメント(クラス)の導出、各クラスと個人属性との関連性の把握を目的とする。推定の結果、宿泊観光旅行回数の需要関数は3クラスとして設定できること、その中でクラス1は旅行非実施クラス(構成比率45%)であり、残りは低頻度クラス(平均旅行回数1.58回、構成比率41%)と高頻度クラス(平均旅行回数5.49回、構成比率13%)に区分することができた。さらに、年収をはじめとして個人属性による旅行回数への影響が高頻度クラスで大きいこと、消費行動、旅行行動への志向が旅行回数に与える影響は年収よりも効果が小さいこと、高頻度クラスになるに従い、高齢者の構成比率が増加するとともに、旅行非実施クラスがいずれの階層にも40%超程度を占めていることが明らかになった。

**Key Word:** tourist behavior, overnight trip, Latent Class Model, tourist preference

## 1. はじめに

本研究は、平成24年度に実施した「国民の観光に関する動向調査」データを用いて、個人の年間宿泊観光旅行回数の潜在クラスモデルによる再現性の検討を行う。分析では、旅行回数を推定する関数を同定するとともに、同一の旅行回数関数を有するクラス(セグメント)別の構成比率の導出、各クラスと個人属性との関連性を明らかにすることを目的とする。

潜在クラスモデルを適用した理由であるが、宿泊旅行回数と個人属性である年収や消費行動・旅行行動への志向との関連では一定の傾向が読み取れるものの、数量化I類モデルなどでは十分な説明力を有するモデルが構築できなかった。これは、観光行動は非日常的活動、非義務的活動であるため、同一デモグラフィック内でのばらつきが大きくなるためと考えられる。これより、異なった旅行回数関数を有するクラスの存在、ならびにそれらに影響を与える個人属性が考えられるため、それらを包括して分析可能な潜在クラスモデルを採用した。

## 2. 本研究の位置づけ

### (1) 国民の観光に関する動向調査<sup>1</sup>について

「国民の観光に関する動向調査」は、(公財)日本観光振興協会(平成23年度以前の日本観光協会)によって、実際に行われた宿泊旅行の実態と今後実施を希望する旅行形態などへの志向を明らかにするために、全国の個人を対象としたアンケート形式によって実施されており、その結果は毎年「観光の実態と志向」として出版されている。その調査項目であるが、前年度に実施した宿泊旅行の回数、目的、同行者、利用交通機関、目的地、利用宿泊施設、旅費等の旅行実績を中心として構成されている。平成24年度調査では、「普段の消費行動(28項目)」、「普段の旅行行動(28項目)」ならびに「行きたい旅行内容(21項目)・同行者の意向」など、消費行動・旅行行動の志向についての設問が加わった。本研究は、平成24年度調査(平成23年の宿泊観光旅行を対象)を用いて分析を行う。

### (2) 既存研究と本研究の位置づけ

個人個人に着目した観光旅行回数に関する研究は、個人属性による宿泊観光旅行回数への影響を分析したもの

<sup>2</sup>, 行動変化を世代, 時代, 年代による影響に分解したもの<sup>3</sup>, 消費行動, 旅行行動への志向についての質問を設け, これらの宿泊観光旅行回数に対する影響を明らかにした研究<sup>4</sup>などがある。これらの研究では, 被験者が同一の旅行回数関数を有しているとしているが, 同じ個人属性でも旅行回数にばらつきを有すると考えられるため, その考慮が必要不可欠といえる。

その対応策として, 潜在クラスモデルの適用が考えられる。潜在クラスモデルの適用例として, 目的地選択において, トライアル層とリピート層で効用関数に差異が存在することを明らかにした研究<sup>5</sup>, 自動車走行量における差異を考慮した研究<sup>6</sup>, まちなか居住促進のための選好クラスを導出した事例<sup>7</sup>がみられるものの, 宿泊観光旅行回数に適用した事例はない。以上より, 本研究の特徴としては, 潜在クラスモデルを宿泊旅行回数の推定に適用し, 同一の旅行回数関数を有するクラスの導出, クラスと個人属性との関連性に着眼する点と考えられる。

### 3. 潜在クラスモデル

行動特性が異なるクラスごとの旅行回数関数を導出することが本研究の目的である。これに照らし合わせて, 既存のモデルを俯瞰すると, 主成分分析(因子分析), クラスタ分析のようなデータからサンプルの特性把握を行う方法があるが, これらの手法は被説明変数との関連性を明らかにすることができない。また, 因子分析における各因子間に斜交性を仮定して, 変数間の構造を把握する手法として, 共分散構造モデルがあるが, 基本的に全サンプルが同一の評価構造を有すると仮定しているため, 異なった評価構造を有するクラスの抽出を行うことはできない。

そこで, 本研究は, 対象集団に観測されない異質性が存在することを前提に, 異質性から導かれるクラス分類とクラスごとの宿泊旅行回数を説明する関数(回帰式)の推定を同時に行うことができる潜在クラス回帰分析を適用した。

クラス数:  $K$ , パラメータ:  $\theta_k$ を所与とし, クラス $k$ に属する確率を  $\pi_k$  ( $\sum_{k=1}^K \pi_k = 1, \pi_k \geq 0$ ) とすると, 個人 $i$ のデータ ( $\mathbf{x}_i, y_i$ ) が観測される確率:  $h_i$ は, 下記のようにあらわすことができる。

$$h(y_i | x_i, \theta_k, \pi_k) = \sum_{k=1}^K \pi_k p(y_i | x_i, \theta_k) \quad \dots(1)式$$

ここで,

- $y_i$ : 個人 $i$ の被説明変数 (年間の宿泊観光旅行回数),
- $\mathbf{x}_i$ : 個人 $i$ の説明変数,
- $p(\cdot)$ : クラスを所与としたときの確率密度関数 (正

規分布を仮定, すなわち

$$p(y_i | x_i, \theta_k) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - \theta_k x_i)^2}{2\sigma^2}\right) \quad \dots(2)式,$$

である。クラス $k$ への所属確率をしめす  $\pi_k$ は, 属性に影響を受けず, 全サンプルで共通とする。さて, (2)式に示したように, 被説明変数:  $y_i$ は, パラメータ:  $\theta_k$ , 説明変数:  $\mathbf{x}_i$ の線形和で示すことができると仮定する。パラメータ:  $\theta_k$ , ならびにクラス $k$ に属する確率:  $\pi_k$ は, (3)式で示される対数尤度の最大化によって導出することができる。

$$\log L = \sum_{i=1}^N \log \sum_{k=1}^K \pi_k p(y_i | x_i, \theta_k) \quad \dots(3)式$$

ここで示したように, パラメータ  $\theta_k$ は個人ごとに導出されるのではなく, 個人が属するクラス $k$ で規定される点に注意する必要がある。

さて, パラメータの導出は対数尤度の最大化によって可能であるが, 最尤法で推定することが困難な場合がある。その理由として, 尤度関数において未知パラメータであるベクトル:  $\theta_k$ , ならびにクラスへの所属確率:  $\pi_k$ の2つの存在があげられる。そこで, EMアルゴリズム (Expectation-Maximization-step) を用いて推定を行うことが多い<sup>8</sup>。まず,  $\pi_k$ は個人 $i$ がクラス $k$ に所属しているとき  $\pi_k=1$ , それ以外の時  $\pi_k=0$ をとる確率変数とする。  $y_i$ と  $\pi_k$ は独立であると仮定し, 適当な初期値を与える。この  $\pi_k$ は, 観測されていないため欠損値であるためデータセットとしては不完全データとよばれ,  $y_i$ と  $\pi_k$ のいずれもわかっている場合は完全データと呼ばれる。EMアルゴリズムでは観測された不完全データをいったん最尤方程式になじみのよい「完全データ」に擬似的に置き換え (E-step), この「擬似的完全データ」を用いてパラメータの擬似的最尤推定値を求め (M-step), さらに得られた推定値から擬似的完全データを作り直し, それを用いて再度パラメータ推定値を求めなおす, という手続きを繰り返す。EMアルゴリズムでは, 繰り返し最適化を行う必要があるため推定時間が長くなるが, 反復ごとに尤度が改善され, 一般的条件の下では収束が保障されている。EMステップを繰り返し行い, 尤度に改善が見られない時点で収束したものと判断して, パラメータ推定が終了する。局所最適解が存在する場合には, 複数の初期値を用いて推定を行うことで問題を解決できるとされている。

この方法は, クラス数を先験的に与えるものであるため, モデルの適合度は赤池情報量基準(AIC), ベイジアン基準量(BIC)により評価される。これらの値が最小になるクラス数を最適クラス数と判断する。

$$AIC = -2\log(\text{最大尤度}) + 2(\text{パラメータ数})$$

$$BIC = -2\log(\text{最大尤度}) + \log(\text{サンプル数}) \cdot (\text{パラメータ数})$$

各クラスのパラメータに選好が反映されていると考えることができるため、推定されたパラメータから各クラスの特徴を事後的に把握する。

#### 4. 被験者の特性と旅行の実態

##### (1) 調査項目と基本集計結果

「第31回国民の観光に関する動向調査」は、平成24年11月にインターネット調査によって実施され、11,146人から回答があった。従って、ネット利用が可能な被験者に限定される点に留意する必要がある。なお、主要調査項目は、下記のとおりである。

- 平成23年に実施した宿泊観光旅行回数、目的、同行者、利用交通機関、利用情報媒体、目的地、利用宿泊施設、旅費等の実績
- 旅行全般に対する意向  
旅行の好き嫌い、主観的旅行頻度、非実施理由
- 個人属性  
性年齢階層、免許保有、家族形態、休日数、世帯所得、居住地域
- 普段の消費行動、普段の旅行行動、行きたい旅行内容

表-1 被験者の個人属性構成比率(1)

性別	運転免許	休日数/月
男性 51%	保有 79%	8日以上(完全週休2日制) 60%
女性 49%	非保有 21%	5日-7日(変則2日制) 28%
-	-	4日以下(変則1日制) 10%
-	-	その他 3%
合計 100%	合計 100%	合計 100%

表-2 被験者の個人属性構成比率(2)

年齢	家族形態	世帯所得(万円)	地域
15-17歳 3%	単身世帯 15%	200未満 11%	北海道 4%
18-19歳 3%	夫婦だけの世帯 27%	200-400未満 24%	東北 7%
20-24歳 5%	夫婦と親の世帯 6%	400-600未満 23%	関東 35%
25-29歳 7%	夫婦と子どもの世帯 35%	600-800未満 13%	甲信越 4%
30-34歳 7%	親と夫婦と子どもの世帯 7%	800-1000未満 8%	中部 13%
35-39歳 9%	その他 10%	1000-1500未満 5%	関西 18%
40-49歳 16%	-	1500-2000未満 1%	中国 6%
50-59歳 14%	-	2000以上 1%	四国 3%
60-69歳 17%	-	未回答 15%	九州 11%
70歳以上 20%	-	-	-
合計 100%	合計 100%	合計 100%	合計 100%

表-1, 2は、被験者の個人属性(7項目)の単純集計結果である。なお、被験者の構成比率は、居住地域毎(北海道~九州・沖縄の8区分)に性年齢階層の構成比率が住民基本台帳と等しくなるようにネット調査会社において設定を行なっている。また、休日数では、農林漁業、自由業を除く就業者の休日数の構成比率を示している。

次に、集計された旅行実態を下記に示す。

- 宿泊観光旅行参加率: 54%
- 宿泊観光旅行平均回数: 1.43回/年・人(表-3)
- 主な旅行目的, 同行者, 代表利用交通機関
- 主に参考にする情報源利用率(上位4位): インターネット(62.4%), ガイドブック(47.3%), パンフレット(37.1%), 家族・友人の話(29.4%)

表-3 被験者の年間宿泊観光旅行回数構成比率

回数	サンプル数	構成比率	累積構成比率
0回	5,086	46%	46%
1回	2,679	24%	70%
2回	1,407	13%	82%
3回	764	7%	89%
4回	445	4%	93%
5回	246	2%	95%
6回	144	1%	97%
7回	105	1%	98%
8回	58	1%	98%
9回	40	0%	98%
10回	33	0%	99%
11回	31	0%	99%
12回	30	0%	99%
13回以上	78	1%	100%
合計	11,146	100%	

##### (2) 消費行動, 旅行行動への志向によるクラスへの導出<sup>9</sup>

消費行動や旅行行動への志向と潜在クラスモデルで求められる宿泊旅行回数との関連を明らかにするために回答データからサンプルを各々10クラスターに分類した。この中で、普段の消費行動は、日常の(旅行以外の)消費行動において心がけていることを明確にするために設定したものであり、消費行動への志向と実旅行行動との関連性把握を念頭にしたものである。また、普段の旅行行動は、特定の旅行ではなく、「旅行全般」に対して自ら好んだり、志向している事項を明確にしながら、実行動との関連性把握を意図した設問である。

まず、普段の消費行動についての28項目の回答データを用いて主成分分析を行った。具体的には、11,146人の28項目の回答「あてはまる:1点」~「あてはまらない:5点」を用いて、バリマックス回転を行いながら累積寄与率が50%を超える第7主成分までを抽出した(表-4)ところ、流行軸, 高級軸, イメージ軸の上位3主成分を抽出できた。なお、累積寄与率は50%強であり、必ずしも高い説明力となっておらず、7つの主成分では消費行動の特徴を十分把握できていないことに留意する必要がある。

次に、導出された個々人の主成分得点を用いて、クラスター分析を行った。分析では主成分得点の大きさをそのまま考慮するためにデータの正規化を行わず、ワード法を採用しながら最終的に10クラスターに分割できると仮定した。表-5, 6は、S1~S10の第1~第7主成分得点の平均値ならびにその特徴を示したものである。

表-4 普段の消費行動の主成分分析結果

主成分	軸・名称	主成分得点説明 (主成分負荷量の大きな小問 番号・丸数字：負値)	寄与率 (%)	累積寄 与率(%)
主成分1	流行軸	流行に敏感であるほど小 (18, 21, 26, 27)	11.7	11.7
主成分2	高級軸	高級志向ほど小 (5, 8, 13, 20)	11.0	22.6
主成分3	イメージ軸	イメージ重視ほど大 (17, 19)	7.3	29.9
主成分4	もったいない軸	もったいないに配慮するほ ど小 (24, 28)	6.8	36.7
主成分5	計画消費軸	計画的なほど大 (2, 4, 10)	6.5	43.2
主成分6	金額節約軸	安い金額を重視するほど、 大 (6, 7, 12)	6.4	49.6
主成分7	環境・利他軸	環境に配慮するほど小 (9, 11)	6.3	55.9

表-5 消費行動クラスター別主成分得点平均値

	第1主成分	第2主成分	第3主成分	第4主成分	第5主成分	第6主成分	第7主成分	評価	構成比率
S1	1.8	-1.5	-1.0	0.2	0.3	-0.2	-1.0	流行に敏感, 高級志向	9%
S2	1.7	0.8	1.5	0.5	-0.6	0.2	-0.4	流行に敏感, イメージ重視	9%
S3	1.2	-1.6	0.2	0.3	0.8	0.1	1.8	流行に敏感, 高級志向, 商品の特色にこだわらない	10%
S4	1.2	2.1	-0.2	-0.8	1.0	0.2	-0.1	非高級志向	11%
S5	0.4	1.5	-1.5	0.1	-1.7	-0.6	0.7	非高級志向, イメージ非重視, 非計画的購買	9%
S6	0.1	-0.7	0.8	-0.9	-0.3	-0.7	-0.4	平均的消費指向	8%
S7	-0.1	0.0	-0.3	1.4	0.1	1.2	-0.1	非コミュニティ購買志向	11%
S8	-1.5	0.0	-0.3	-1.2	-0.4	0.1	0.3	流行に敏感, コミュニティ購買志向	13%
S9	-1.9	1.5	0.3	0.8	0.6	-0.1	-0.3	流行に敏感, 非高級志向	11%
S10	-2.2	-1.8	0.3	0.2	0.0	0.0	-0.2	流行に敏感, 高級志向	10%

表-6 消費行動クラスターの特徴

説明 (主要な性年齢層)	特徴
S1 流行に無頓着だが高級品を好むタイプ (男女60代以上)	新製品やロコミをあまりチェックせず, お金を払っても, ワンランク上のサービスを求める
S2 流行に無頓着だが, マス情報に左右されやすいタイプ (すべての性年齢層, 若干女性層が多い)	新製品やロコミをあまりチェックせず, テレビCMやパッケージにつられて購買してしまう
S3 流行に無頓着で高級品を好むが中身には関心がないタイプ (男女20代)	新製品やロコミをあまりチェックせず, お金を払っても, ワンランク上のサービスを求める。また, 環境に配慮した商品や日本製・日本産など商品の特徴はあまり考慮しない
S4 まず価格を重視するタイプ (すべての性年齢層, やや女性層が多い)	気に入っても, 価格が高ければ購入しない
S5 価格重視で衝動買いするタイプ (男性50代以上)	気に入っても, 価格が高ければ購入せず, テレビCMやパッケージには影響を受けない。買ったことがない銘柄でもあまり検討しない
S6 はっきりとした傾向が見られないタイプ (すべての性年齢層)	これといった特徴がない層
S7 社会とのつながりを重視しないタイプ (すべての性年齢層, 若干男性40代以上が多い)	バザーでの購入や友人等からお下がりを貰わない
S8 流行に敏感で社会とのつながりを活かすタイプ (男女30代以下)	新製品やロコミを常にチェックしながら, バザーで購入したり友人等からお下がりをもらう
S9 流行に敏感だが価格にもシビアなタイプ (すべての性年齢層)	新製品やロコミを常にチェックしながら, 気に入っても, 価格が高ければ購入しない
S10 流行に敏感で高級品を好むタイプ (30代以下, 若干男性10,20代が多い)	新製品やロコミを常にチェックしながら, お金を払っても, ワンランク上のサービスを求める

表-7 各性年齢階層の消費行動クラスター別構成比率

セグメント	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
男性・10-17歳	3%	4%	15%	4%	7%	10%	3%	21%	7%	25%
男性・18-19歳	3%	5%	20%	3%	8%	7%	6%	14%	18%	21%
男性・20-24歳	5%	3%	24%	4%	4%	7%	7%	18%	10%	17%
男性・25-29歳	4%	4%	21%	3%	4%	7%	6%	20%	18%	18%
男性・30-34歳	6%	5%	18%	5%	5%	6%	9%	16%	11%	18%
男性・35-39歳	5%	7%	15%	7%	7%	6%	10%	19%	9%	15%
男性・40-49歳	8%	8%	12%	9%	6%	7%	12%	13%	9%	18%
男性・50-59歳	11%	10%	8%	9%	10%	9%	12%	11%	8%	11%
男性・60-69歳	17%	10%	6%	10%	18%	9%	11%	10%	18%	5%
男性・70歳-	19%	11%	2%	9%	15%	7%	11%	9%	9%	7%
男性合計	11%	8%	10%	8%	10%	8%	10%	14%	9%	12%
女性・10-17歳	1%	8%	11%	9%	8%	10%	3%	24%	12%	14%
女性・18-19歳	2%	9%	17%	9%	5%	9%	9%	16%	11%	18%
女性・20-24歳	4%	8%	27%	8%	5%	6%	8%	15%	9%	10%
女性・25-29歳	4%	8%	18%	10%	3%	10%	5%	17%	18%	13%
女性・30-34歳	5%	18%	16%	11%	4%	8%	5%	14%	10%	14%
女性・35-39歳	6%	11%	9%	14%	4%	9%	9%	14%	11%	14%
女性・40-49歳	9%	12%	16%	12%	3%	10%	9%	12%	16%	11%
女性・50-59歳	9%	18%	14%	13%	5%	18%	7%	9%	17%	10%
女性・60-69歳	19%	14%	3%	13%	7%	18%	7%	8%	19%	8%
女性・70歳-	21%	18%	2%	12%	8%	12%	6%	9%	9%	8%
女性合計	11%	12%	8%	12%	5%	11%	7%	12%	12%	11%
合計	11%	10%	9%	10%	8%	9%	9%	13%	11%	11%

表-7は, 上記クラスターごとの性年齢階層別構成比率を示したものである。表-7より, 流行に無頓着であるS1, S2は高齢者層でみられる一方, 流行に敏感であるS8, S9, S10は40歳代以下で構成比率が大きいことがわかる。また, 男女で構成比率が異なるS2(女性:多), S4(女性:多), S7(男性:多)や, 20-24歳代で構成比率が高いS3クラスターもある。以上より, 概ね消費行動によるクラスターは, 性年齢階層との関連性を見出すことができ, 個人の性年齢階層が, 消費行動と密接に関連する余暇時間や可処分所得, ライフスタイルを規定する大きな影響要因と考えられる。

これらと同様に, 「普段の旅行行動」データを用いて分析を行った。表-8は, 累積寄与率が50%を超える第6主成分まで抽出した結果である。さらに, 導出された個々人の主成分得点を用いて, クラスター分析を行った。「旅行行動」でも10クラスターに分割することができる。消費行動と同様に, 被験者の志向が多様であることから説明力が十分高いとは言えない。

表-8 普段の旅行行動の主成分分析結果

主成分	軸・名称	主成分得点説明 (主成分負荷量の大きな小問 番号・丸数字：負値)	寄与率 (%)	累積寄 与率(%)
主成分1	情報取得軸	普段から情報取得するほど大 (21, 22, 23, 24, 26)	14.3	14.3
主成分2	計画立案軸	しっかり計画を立てるほど大 (2, 6, 7)	10.7	25.0
主成分3	マスワザム(イメージ・お手軽)軸	ポスター, TVに影響を受け, パック旅行や有料ガイドを頼むほど小 (20)	8.0	33.0
主成分4	節約消費	節約消費するほど大 (9, 12, 15)	7.2	40.2
主成分5	こだわり軸	場所などをこだわるほど小 (10, 14)	7.2	47.4
主成分6	地域社会軸	交流志向や公共交通機関利用意向ほど大 (13, 18)	6.2	53.6

表-9 旅行行動クラスター別主成分得点平均値

	第1主成分	第2主成分	第3主成分	第4主成分	第5主成分	第6主成分	評価	構成比率
R1	-2.7	-0.1	-1.3	0.1	0.1	0.1	非情報取得, イメージ・お手軽な旅行志向	11%
R2	-1.5	2.2	0.4	0.1	0.0	-1.1	非情報取得, 旅行計画立案	9%
R3	-1.4	-0.3	1.5	-0.1	-1.6	1.1	非情報取得, イメージ・お手軽な旅行を非志向	9%
R4	-0.7	-0.1	1.1	1.4	1.5	0.2	節約志向, 非こだわり志向	8%
R5	-0.5	-3.0	0.9	-0.7	0.9	-0.7	非旅行計画立案	8%
R6	-0.2	1.0	-0.2	-1.9	0.1	0.4	非節約志向	9%
R7	0.0	-1.0	-0.7	0.5	-1.1	-0.7	平均的(非計画立案, こだわり志向)	10%
R8	0.4	0.9	-0.7	0.5	0.8	0.5	平均的(計画立案, こだわり小)	8%
R9	2.3	-0.7	-1.2	0.1	-0.1	0.4	情報取得し, イメージ・お手軽な旅行志向	16%
R10	2.4	1.1	1.3	0.1	-0.2	-0.4	情報取得, イメージ・お手軽な旅行を非志向	12%

表-9, 10は, R1~R10の第1~第6主成分得点の平均値ならびに特徴を示したものである。さらに, 表-11は, クラスター別の性年齢階層別構成比率を示したものである。表-11より, 若年層でR3の構成比率が大きくなっていること, R8で女性20-30歳代の構成比率が大きいことが特徴としてみられる。一方, 消費行動と比較すると性年齢

階層別構成比率との関連性が弱いと考えられる。これから、旅行行動の志向は、性年齢階層といった個人属性によって強く規定されないものと推察できる。

表-10 旅行行動クラスターの特徴

グループ名称 (主要な性年齢階層)	特徴
R1 情報は取得せず、マックス要素：大 (全ての性年齢階層)	普段は旅行情報を取得せず、ポスター、TVに影響を受け、バック旅行や有料ガイドを頼む。
R2 情報は取得しないが、旅行計画は立案する (全ての性年齢階層、若干女性層が多い)	普段は旅行情報を取得せず、事前に旅行計画を立案。
R3 情報は取得しないが、マックス要素も小さい (全ての性年齢階層、若干20代以下が多い)	普段は旅行情報を取得せず、ポスター、TVに影響を受けず、バック旅行も選ばない傾向。
R4 節約消費を心がけ、こだわりも小さい (全ての性年齢階層)	節約消費に心がけ、目的地や宿泊施設などにこだわらない。
R5 旅行計画を立てない (全ての性年齢階層)	事前に旅行計画を立案しない傾向
R6 非節約消費志向 (全ての性年齢階層、若干60代以上：多)	節約消費を心がけない層
R7 平均的 (全ての性年齢階層)	平均的旅行指向であるが、事前に旅行計画を立案せず、目的地や宿泊施設に若干のこだわりがある。
R8 平均的 (全ての性年齢階層、若干20、30代が多い)	平均的旅行指向であるが、事前に旅行計画を立案して、目的地や宿泊施設のこだわりは小さい。
R9 情報取得し、マックス要素：大 (全ての性年齢階層、若干男性が多い)	普段から旅行情報を取得し、ポスター、TVに影響を受けバック旅行を選択。
R10 情報取得し、マックス要素が小さい (全ての性年齢階層、若干女性が多い)	普段から旅行情報を取得し、ポスター、TVに影響を受けずバック旅行も選ばない傾向。

表-11 各性年齢階層の旅行行動クラスター別構成比率

セグメント	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
男性・10-17歳	5%	6%	36%	7%	6%	7%	7%	5%	12%	8%
男性・18-19歳	4%	8%	26%	7%	8%	4%	10%	11%	12%	10%
男性・20-24歳	6%	8%	26%	10%	4%	5%	12%	10%	10%	8%
男性・25-29歳	5%	8%	27%	10%	9%	5%	9%	12%	7%	8%
男性・30-34歳	7%	8%	23%	12%	7%	4%	9%	12%	8%	10%
男性・35-39歳	8%	5%	17%	14%	10%	5%	12%	9%	10%	9%
男性・40-49歳	6%	8%	17%	13%	11%	6%	11%	9%	10%	9%
男性・50-59歳	8%	8%	15%	13%	11%	7%	7%	8%	10%	12%
男性・60-69歳	8%	12%	11%	10%	10%	10%	9%	7%	11%	11%
男性・70歳-	11%	13%	11%	11%	10%	13%	7%	6%	10%	7%
男性合計	8%	9%	17%	12%	10%	8%	9%	8%	10%	9%
女性・10-17歳	8%	8%	32%	6%	7%	4%	8%	7%	11%	9%
女性・18-19歳	14%	9%	22%	10%	5%	10%	8%	9%	8%	5%
女性・20-24歳	11%	13%	15%	8%	7%	6%	11%	13%	7%	8%
女性・25-29歳	9%	11%	18%	13%	5%	5%	10%	17%	6%	7%
女性・30-34歳	9%	9%	16%	14%	6%	6%	6%	14%	6%	13%
女性・35-39歳	7%	10%	17%	14%	9%	4%	8%	15%	7%	11%
女性・40-49歳	8%	12%	12%	17%	9%	7%	7%	11%	7%	11%
女性・50-59歳	8%	13%	12%	15%	9%	9%	7%	9%	6%	12%
女性・60-69歳	8%	17%	14%	12%	9%	13%	4%	5%	5%	12%
女性・70歳-	9%	15%	14%	11%	11%	19%	4%	5%	6%	9%
女性合計	8%	13%	15%	13%	8%	10%	6%	10%	6%	10%
合計	8%	11%	16%	12%	9%	9%	8%	9%	8%	10%

クラス数4を超えると収束せず、パラメータを導出することができなかった。これは、表-3に示すように旅行回数が多くなるに従って構成比率が低減し、ピークが見られないことから適当なクラスを一意に導出できないためと考えられる。BIC, AICは、「限られたデータのもとで現実を最もよく把握するためにはデータ制約に応じた単純なモデルが必要とされる」という「ケチの原理」に基づいており、より小さいモデルほど好ましいと判断される<sup>10</sup>。BICはAICよりもサンプル数が考慮されるため、よりシンプルなモデルが評価されることになる。これより、クラス数3 (表中表記は3(1)) が適当と判断できる。

表-12 クラス数別モデル推定結果

クラス数	1	3(1)			3(2)			
構成比率	100%	45%	41%	13%	45%	41%	13%	
平均回数	1.41	0.00	1.58	5.49	0.00	1.58	5.25	
定数項	1.08	0.00	1.50	5.46	0.00	1.55	5.43	
性別	女性	0.15	0.00	0.03	0.08	0.00	0.03	0.02
世帯構成	夫婦・子	-0.34	0.00	-0.15	-0.50	0.00	-0.15	-0.54
	親・夫婦・子	-0.48	0.00	-0.20	-1.36	0.00	-0.19	-1.44
年収	600~799万円	0.46	0.00	0.14	0.34	0.00	0.13	0.31
	800~999万円	0.80	0.00	0.21	1.24	0.00	0.20	1.17
	1000~1499万円	0.84	0.00	0.19	1.24	0.00	0.16	1.02
	1500~1999万円	1.74	0.00	0.46	4.35	0.00	0.43	4.26
	2000万円以上	2.27	0.00	0.87	3.69	0.00	0.94	3.70
免許保有		0.24	0.00	0.10	-0.20	0.00	0.11	-0.16
休暇・月8日以上		0.14	0.00	-0.02	0.61	0.00	-0.02	0.54
消費志向・S8							-0.15	0.11
消費志向・S9							-0.12	-0.84
旅行志向・R2							-0.18	-1.34
旅行志向・R3							0.03	0.95
旅行志向・R4							0.23	1.04
旅行志向・R7							-0.15	-0.55
旅行志向・R8							-0.13	-0.79
旅行志向・R9							-0.16	-0.81

斜体:  $|t\text{値}| \leq 1.96$  の場合. 係数=0.0: 灰色, 負値: 赤で表記.

表-13 クラス数別モデル評価指標

クラス数	BIC	AIC	決定係数
1	51,414	51,238	0.05
2	39,919	39,429	0.50
<b>3(1)</b>	<b>8,301</b>	<b>7,906</b>	<b>0.65</b>
<b>3(2)</b>	<b>8,391</b>	<b>7,820</b>	<b>0.66</b>
4	-	-	-

## 5. 潜在クラスモデルの推定

潜在クラスモデルの推定では、クラス数設定と説明変数の選択の2つを検討する必要がある。本研究では、基本的なモデルであるクラス数1の場合にt値が有意な変数を確認し、その変数を用いながらクラス数を変化させてBIC, AICがより小さくなるクラス数を決定した。なお、クラス数が1の場合の推定結果は、通常の数値化I類と同値となる。

まず、個人々人を対象とする11,146人のデータを用いてクラス数1の推定を行った。説明変数は個人属性データであるため、該当時:1, 非該当時:0とする質的データに変換している。表-12は、その結果であるが、5%有意な変数として性別、世帯構成、年収、免許保有が抽出できた。年齢階層では若年層で、居住地域では多くの地域で有意とならなかったため、これらを推定から除外した。

そして、クラス数1と同様の説明変数を用いてクラス数を変化させたところ、表-13のようになった。この時、

次に、決定係数であるが、被説明変数の全平方和に対する推定値の回帰平方和の比率、もしくは全分散から誤差平方和を減じたものの比率として導出できる。さて、複数クラスの場合は、それぞれのクラスについて推定値がある。今回は実績の旅行回数と最も残差が最小なものとの組み合わせの中で算出した決定係数とした。この場合、クラス数1の0.05に比べてクラス数2の場合:0.50、クラス数3の場合:0.65と大きく向上していることがわかる。旅行回数は0回が46%を占め、年数回程度の低頻度であるため決定係数が低くなるが、クラスを複数設定することにより改善できるものと考えられる。以上の考え

に基づき、クラス数3が最も説明力が高いと判断された。

この時、パラメータの符号条件、大小関係を見ると、これまでの研究と同様の傾向を示し、t値も有意となっている。この中でクラス1はすべてのパラメータが0.0となっているため旅行非実施クラスを再現し、構成比率は45%となった。これに対して、クラス2は低頻度クラス（平均旅行回数1.58回、構成比率41%）を、クラス3は高頻度クラス（平均旅行回数5.49回、構成比率13%）を再現している。両者のパラメータは、定数項の他、年収をはじめとしてクラス3の方が値が大きく、旅行回数への影響が大きいことがわかる。

これらの変数に加えて、消費行動、旅行行動への志向が旅行回数に与える影響を明らかにするために、志向に関する変数を追加して推定を行った。ここでは、4.2節で導出した消費、旅行行動のクラスターを用いて、それぞれをダミー変数として導入した。最終的に有意な変数のみに限定したものが、表-12、表-13におけるクラス数3(2)のモデルである。個々のパラメータが有意であり、クラスターによって回数の増加、減少傾向を導くことができた。文献4では年収と同程度の影響とされていたが、クラスを分けた本研究ではそれよりも効果が小さいことがわかる。また、多くのクラスターで高頻度クラスの方が影響が大きくなっており、志向の影響がより大きくなると考えることができる。モデル全体の説明力自体は大きく改善されていないもののデモグラフィック変数以外の影響を明示できたと考えられる。

さて、これら3クラス各々を構成するサンプルはどのような属性を有しているのであろうか。推定値と実績値の残差が最小となるクラスを各サンプルの所属クラスと判別し、同一属性毎にクラス構成割合を集計することにより、その傾向を把握できる。集計する属性単位は推定に利用していないものが望ましいため、年齢階層をここでは用いる。表-14は年齢階層別所属クラス構成比率である。クラスが高くなるに従い、高齢者の構成比率が増加するとともに、旅行非実施クラスがいずれの階層にも40%超程度を占めていることが明らかになった。

表-14 年齢階層別所属クラス構成比率

年齢階層	クラス1	クラス2	クラス3
19歳以下	55%	40%	5%
20-24歳	43%	47%	10%
25-29歳	46%	42%	12%
30-34歳	47%	43%	9%
35-39歳	42%	49%	9%
40-49歳	49%	41%	10%
50-59歳	50%	38%	12%
60-69歳	42%	40%	18%
70歳以上	41%	40%	19%

## 5. まとめ

本研究は、個人の年間宿泊観光旅行回数の潜在クラスモデルによる再現性の検討を行った。分析では、旅行回数を推定する関数を同定するとともに、同一の旅行回数関数を有するクラス(セグメント)別の構成比率の導出、各クラスと個人属性との関連性を明らかにした。

推定の結果、宿泊観光旅行回数の需要関数は3クラスとして設定できること、その中でクラス1は旅行非実施クラスであり、残りは低頻度クラス（平均旅行回数1.58回、構成比率41%）と高頻度クラス（平均旅行回数5.49回、構成比率13%）に区分することができた。さらに、年収をはじめとして個人属性による旅行回数への影響が高頻度クラスで大きいこと、消費行動、旅行行動への志向が旅行回数に与える影響は年収よりも効果が小さいこと、高頻度クラスになるのに従い、高齢者の構成比率が増加するとともに、旅行非実施クラスがいずれの階層にも40%超程度を占めていることが明らかになった。

今後の課題として、説明力向上に向けたクラス数、変数設定に関する考察、今後の宿泊旅行総量への知見の蓄積があげられる。なお、本研究はJSPS科研費26360077の助成を受けたものである。

## 参考文献

- 1) 日本観光振興協会 HP(<http://www.nihon-kankou.or.jp/home/jigyou/public.html>), 2013. 06. 26閲覧
- 2) 古屋秀樹, 兵藤哲朗, 森地茂: 発生回数の分布に着目した観光交通行動に関する基礎的研究, 第28回都市計画学会学術研究論文集, pp. 319-326, 1993
- 3) 日比野直彦, 森地茂: 世代の特徴に着目した国内観光行動の時系列分析, 土木計画学研究・論文集, Vol. 23, No. 2, pp. 399-406, 2006
- 4) 古屋秀樹, 全相鎮: 旅行者の志向を考慮した宿泊観光旅行の基礎的分析, 第48回土木計画学研究発表会講演集, No. 118 (CD-ROM), 2013
- 5) 小李克行, 西村雅史, 岩倉成志: 観光需要におけるトライアル層とリピート層の嗜好性に関する考察, J-Rail2005, pp. 407-410, 2006
- 6) 谷下雅義: 世帯の多様性を考慮した人口密度が自動車走行量に及ぼす影響分析-オーナーインタビュー調査を用いた潜在クラス回帰分析-, 第42回土木計画学研究発表会・講演集, No. 176 (CD-ROM), 2011
- 7) 溝上章志, 藤見俊夫, 内添啓太: まちなか居住促進のための選好セグメントの分離とその特性分析, 土木学会論文集D3, Vol. 69, No. 2, pp. 121-134, 2013
- 8) 里村卓也: マーケティング・モデル(第6章 セグメンテーションと潜在クラスモデル), 共立出版, 2010
- 9) 前掲4
- 10) 刈屋武昭: 計量経済分析の基礎と応用, 東洋経済新報社

(2014. 8. 1受付)