

# ベイズ・マルチレベルモデルを用いた観光交流人口に関する基礎的研究\*

## Analysis on Tourism Demand via Bayesian Multilevel-model\*

古谷知之\*\*・大田ひとみ\*\*\*

By Tomoyuki FURUTANI\*\*・Hitomi Ota\*\*\*

### 1. はじめに

我が国の観光政策では、「住んでよし、訪れてよし」という、定住人口と交流人口双方を対象とした政策に力点が置かれている。とりわけ定住人口数の減少が予想される地域では、観光客や中長期滞在者などの交流人口数の増加を目的とした政策が、税収増など地域経済などの活性化に繋がると期待されている。

観光政策を立案する際には、しばしば、国内外の観光市場における観光交流人口ODデータなどを用いて、マーケットシェアに基づくポートフォリオ・マトリックス分析や目的地選択モデルの構築などを行い、デスティネーション・マーケティングを展開する。しかし、観光交流人口ODデータを用いてマーケットシェアモデルや目的地選択モデルを構築する際、選択枝数が多いことや、モデル構築を目的にデータ収集されていないため<sup>(例は註1)</sup>、現況再現性の高いモデルが推定できないことが多い。

近年、人の移動に関する分析にも、マルチレベルモデルを用いて、個人や地域、集団間の異質性を明示的に考慮した分析手法が適用されつつある<sup>2)</sup>。また、最尤推定法により解が得られないような場合に、ベイズ推定によりマルチレベルモデルの未知パラメータを推定する方法も用いられている<sup>(例は註3)</sup>。観光市場分析におけるマーケットシェアモデルや目的地選択モデルの推定に際し、ベイズ・マルチレベルモデルを適用することで、現況再現性の高いモデルを構築する可能性があると考えられるが、従来このような研究は行われていない。

そこで本研究では、観光需要分析におけるマーケットシェアモデル或いは目的地選択モデルの構築を前提に、ベイズ・マルチレベルモデルによるロジットモデルを提案し、その有効性を示すことを目的とする。

第二章では研究の方針と方法を述べ、第三章ではモデルの適用結果を示す。最後に研究の課題と今後の展望についてまとめる。

\*キーワード：観光、ベイズ統計、マルチレベルモデル

\*\*正員、博（工）、慶應義塾大学総合政策学部

(神奈川県藤沢市遠藤5322、

TEL:0466-49-3620、E-mail:maunz@sfc.keio.ac.jp)

\*\*\*日本マクドナルドHD株式会社

### 2. 研究の方針と方法

#### (1) 研究の方針

本研究では、都道府県間の観光ODデータを用いたマーケットシェア分析或いは目的地選択モデルを前提として、マルチレベルモデルに基づく集計型ロジットモデルを構築する。とりわけ、ODペア毎の異質性を考慮して、モデルパラメータを推定する。ODペア毎にモデルパラメータを推定する場合、最尤推定法では解が得られないと考えられるため、階層ベイズ法によるシミュレーション（マルコフ連鎖モンテカルロ法：MCMC）を適用する。

観光需要の分析には、第4回全国幹線旅客純流動調査（2005年）<sup>4)</sup>、観光地の地域属性として「地域ブランド戦略サーベイ」<sup>5)</sup>に記載されている地域ブランド力や魅力度などの指標を用いる。

#### (2) 分析の方法

##### a) モデルの概要

発地*i*に対して着地*j*を選択する場合のマーケットシェア  $p_{ij}$  を説明する（目的地選択）モデルを以下のように定式化する<sup>6)</sup>。発地*i*に対して着地*j*を選択する場合の効用関数  $U_{ij}$  を、次式のようにあらわす。

$$U_{ij} = X_{ij}\beta_{ij} + \varepsilon_{ij}, \\ = V_{ij} + \varepsilon_{ij},$$

$$\varepsilon_{ij} \sim \text{i.i.d.}\mathcal{N}(0, \sigma_{ij}^2)$$

地域数が  $n$ 、説明変数が  $k$  個のとき、 $\beta_{ij}$  は  $ij$  地点間の

組み合わせに対する未知パラメータベクトル、 $X_{ij}$  は説

明変数行列、 $\varepsilon_{ij}$  は誤差項を意味する。効用関数の確定

項  $V_{ij}$  を用いて、発地*i*に対する着地*j*のマーケットシ

ェア  $p_{ij}$  は、次式のようにあらわされる。

$$p_{ij} = \frac{\exp(V_{ij})}{\sum_j \exp(V_{ij})}$$

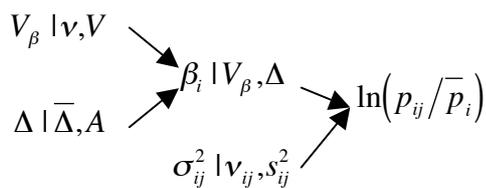
ここで未知パラメータ  $\beta_{ij}$  は、混合効果を示す確定項  $z_{ij}$  と誤差項  $v_{ij}$  とに分解できる。超パラメータ  $v$ 、 $V$ 、 $A$ 、 $\Delta$  を階層事前情報として用いることにより、誤差項の分散  $\sigma_{ij}^2$ 、 $V_\beta^2$  を事後情報として得ることが出来る<sup>3)</sup>。ただし、 $s_{ij}$  は誤差項  $\varepsilon_{ij}$  の分散  $\sigma_{ij}^2$  に対するスケールパラメータ、 $IW(\cdot)$  は逆ウィシャート分布である。

$$\begin{aligned} \beta_{ij} &= \Delta' z_{ij} + v_{ij}, \\ \Delta &= [\delta_1, \dots, \delta_k], \\ \sigma_{ij}^2 &\sim v_{ij} s_{ij}^2 / \chi_{v_{ij}}^2, \\ v_{ij} &\sim \text{i.i.d.} N(0, V_\beta^2), \\ V_\beta &\sim IW(v, V), \end{aligned}$$

$$\text{vec}(\Delta) | V_\beta \sim N(\text{vec}(\bar{\Delta}), V_\beta \otimes A^{-1})$$

モデルパラメータを推定する際には、発地  $i$  に対するマーケットシェアの平均値  $\bar{p}_i$  および説明変数の平均値  $\bar{X}_i$  を基準として対数変換する。

すなわち、以下のようにモデルパラメータを階層ベイズ推定する。



#### b) モデル推定に用いる変数

観光目的来訪者が訪問先を選択する場合、発着地間のアクセシビリティ、観光地側の観光関連施設やサービス、受け入れ体制の整備状況、実際の整備水準よりも来訪者が観光地に対して感じる魅力やブランド力、来訪意向等に影響を受けると考えられる。他方、観光地側が自地域へのデスティネーションマーケティングを展開する際には、発地となる地域の居住者に地域の魅力情報を発信することが少なくない。

本研究では、第4回全国幹線旅客純流動調査（2005年）<sup>3)</sup>の観光目的ODデータ（代表交通手段＝全機関、発地＝居住地）及び地域間移動時間データを用いる。幹線

旅客純流動調査では、都道府県間を超える旅客流動を対象としており、首都圏、中京圏、近畿圏における大都市圏内の流動は都道府県内々の日常流動として対象外となっているが、都道府県単位での観光旅客流動に基づくマーケットシェアを分析するには、十分と考えられる。分析を単純化するため、地域間移動時間については、代表交通機関別移動時間のうち最小の移動時間となるものを採用した。地域魅力度に関する指標は、「2008地域ブランド戦略サーベイ」に記載されている、2006年調査データのうち、地域ブランド力、訪問意向、居住意向に関する地域ブランド力評価値、土産品、郷土料理、観光施設、自然資源、イベント、歴史資源、文化施設、宿泊施設、インフラ（施設・道路）整備状況に関する地域ブランド魅力度評価値のデータを用いることを検討した。変数間の多重共線性などの影響を除外し、最終的に、地域ブランド力、自然資源及びインフラ整備に関する魅力度指標を説明変数として採用した。

#### c) モデルパラメータのベイズ推定

本稿では、MCMCの代表的な手法であるギブズ・サンプラーを用いてモデルパラメータをベイズ推定した。MCMCは3,000回設定し、MCMCの収束判定を行った上で、稼働検査期間を1,000回とした。

### 3. モデルの推定結果

#### (1) モデルパラメータの推定結果

##### a) 変数別にみた推定結果

稼働検査期間を除くモデルパラメータの推定結果を、変数毎に基本統計量をまとめたものを表-1に示す。平均値を見ると、所要時間と定数項は負であり、他の変数は正となっている。標準偏差がいずれも小さく、図-1にも示されているように、モデルパラメータが平均の周辺に分布していることから、モデルパラメータの推定結果として妥当と言える。平均値の符号と2.5%値または97.5%値の符号と異なる変数があるため、ODペア毎に変数を取捨選択することによりモデル全体を改善できる可能性があるが、その点については今後の課題としたい。

なお、提案したモデルは、最尤推定法では解が得られなかった。

表-1 パラメータの推定結果(全体)

変数名	平均値	2.5%	97.5%	s. d.
定数項	-2.245	-3.570	-0.914	0.68
所要時間	-0.036	-0.118	0.048	0.04
ブランド力	0.008	-0.166	0.183	0.08
自然資源	0.041	-0.428	0.514	0.24
インフラ整備	0.100	-0.965	1.139	0.55

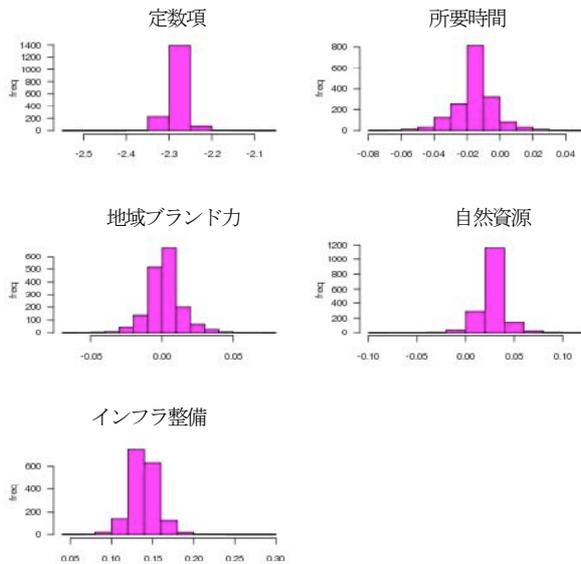


図-1 パラメータの分布

b) 都道府県別の推定結果

モデルパラメータの推定結果の平均値をもちいて、都道府県別のパラメータ推定結果について示してみよう。

例えば、地域ブランド魅力度指標として自然資源を取り上げた場合、東京都を発地とするモデルパラメータの分布は、図-2のようになる。この結果から、青森県や宮城県、長野県などではモデルパラメータが正で相対的に絶対値が大きいことから、これらの地域では自然資源に関する地域ブランド魅力度を高めることにより、東京都からの観光客のシェア増加につながる可能性があるといえよう。

インフラ整備指標に関するパラメータ分布を見ると(図-3)、山形県、長野県、愛知県、島根県などの自治体でパラメータの値が正で相対的に大きいことがわかる。これらの自治体では、自治体内でのインフラ整備に関するブランド力を高めることにより、東京都からの観光客誘致に繋がる可能性があることがあるといえる。山形県や島根県は、インフラ整備に関する地域ブランド魅力度が、全国的に見て相対的に非常に小さい結果が示されているが、地域のインフラ整備を進めることで、東京都からの観光客誘致が期待できる可能性があると考えられる。

長崎県を着地とする所要時間のパラメータ分布をみると、図-4のようになる。この結果から、長崎県に対する所要時間改善による観光客数増加につながる可能性がある都道府県は、福岡県や佐賀県などの近県に加え、愛媛県や青森県などがあげられることが示された。

以上のように、ODペア毎にモデルパラメータを推定することで、自治体が他の地域に対してどのようなデステーションマーケティングを展開すればよいかを、個別に検討することが出来るようになる。

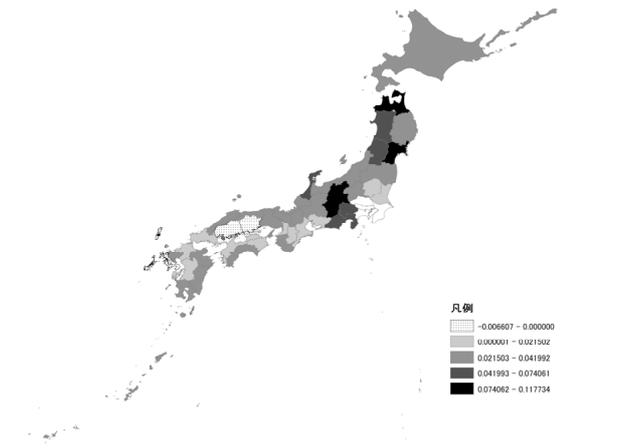


図-2 「自然資源」のパラメータ分布(平均、発地=東京都)

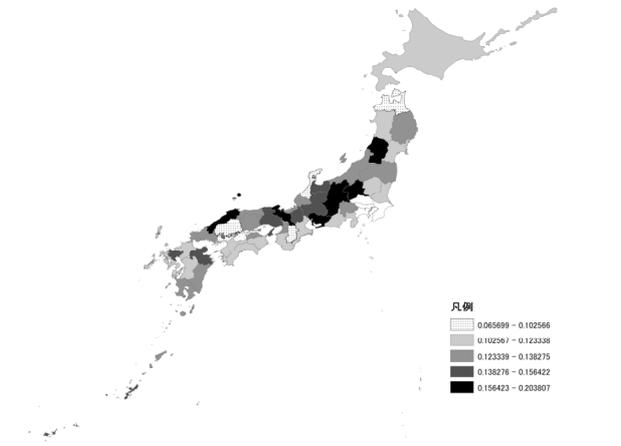


図-3 「インフラ整備」のパラメータ分布(平均、発地=東京都)

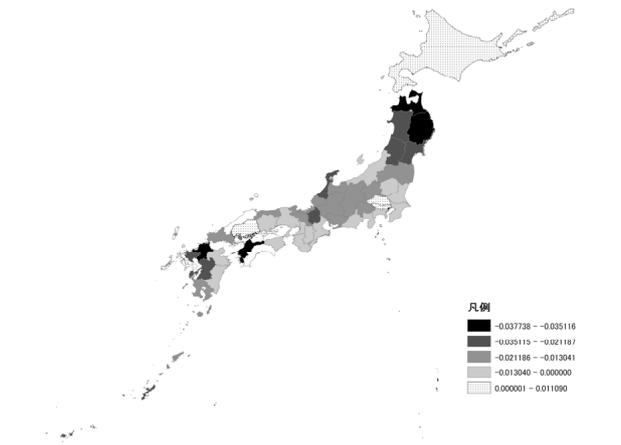


図-4 「所要時間」のパラメータ分布(平均、着地=長崎県)

c) 現況再現性

得られたパラメータ推定結果の平均値を用いてマーケットシェアモデルを構築し、都道府県別観光客発生交通量をかけることにより、観光目的ODデータを再現した。その結果、OD交通量ゼロのODペアを除き、現況観光ODデータと再現した観光ODデータの相関は図-5のとおりであり、R<sup>2</sup>値は0.9993、RMSE誤差は73.5であった。

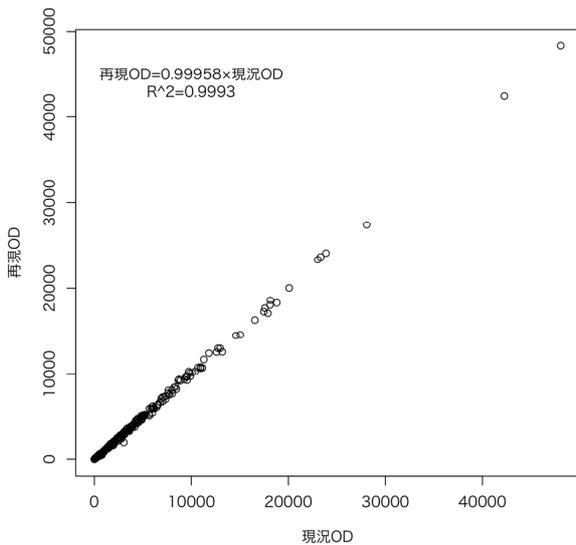


図-5 モデルの現況再現性

表-2 RMSE誤差の比較

モデル	RMSE誤差
提案モデル	73.5
モデル1	2767.9
モデル2	4298.8

モデル推定結果の現況再現性の高さを比較検討するため、以下の2つのモデルも比較し、RMSE誤差を算出した。モデル1は、ODペア毎の異質性を考慮しない従来型のモデルをベイズ推定した。モデル2は、提案したモデルの異質性に関する仮定を緩和して着地毎にモデルパラメータをベイズ推定したモデルである。

**モデル1**：ODペア間の異質性を考慮しない

**モデル2**：着地毎の異質性を考慮

この結果、RMSE誤差は表-2のようになり、提案したモデルが最もRMSE誤差が小さい結果となった。モデル1はR<sup>2</sup>値が0.14であった。モデル2でRMSE誤差が大きいのは、沖縄を着地とする観光客数が過大推計されたためである。

以上のような結果から、提案したモデルでの現況再現性は非常に高いといえる。

#### 4. おわりに

本研究では、ベイズ・マルチレベルモデルと集計型ロジットモデルを組み合わせた、マーケットシェアモデルを提案した。提案したモデルでは、誤差項の分散にODペア間の異質性を考慮することで、誤差項の分散不均一を仮定したモデルとなっている。

全国幹線旅客流動調査の観光目的ODデータなどを用いて適用した結果、ODペア毎にモデルパラメータをベイズ

推定することにより、非常に現況再現性の高いモデルを構築することが出来た。それにより地域ブランド力や観光資源、社会資本整備に関するイメージ、魅力度の改善が観光客の増加に与える影響を、自治体の組み合わせ毎に考察することができた。ブランド力や魅力度に関する指標は、観光関連の社会基盤整備による影響を直接計測できる指標ではないが、観光行動に直接影響を与える指標と考えられる。今後、ブランドやイメージの改善、地域のソフトパワーの向上が、観光集客交流人口の増減、及びそれに伴う自治体の税収増や雇用創出効果に与える影響を検討していきたい。

ベイズ推定は、サンプル数が少ない場合や最尤推定法などで解が得られない場合のモデル推定に有効であると言われているが、本研究で提案したようなモデルでも、モデルに併せてパラメータを推定できるベイズ推定の有効性が示されたと言える。今回は、マーケットシェアモデルを例に分析を行ったが、集計型の目的地選択モデルや手段選択モデルなど、観光行動の分析への適応も、有益と考えられる。

ODペア毎にモデルパラメータを推定するような場合には、ODペア毎に変数を取捨選択することで、モデルの改善が期待できるが、この点は今後の課題としたい。

#### 参考文献

- 1) 味水佑毅：観光統計の整備における「活用の視点」の重要性，国際交通安全学会誌，Vol. 31, No. 3, 236-245.
- 2) Bottai, M., Salvati, N., Orsini, N. : Multilevel models for analyzing people's daily movement behavior, Journal of Geographical System, vol. 8, pp. 97-108, 2006.
- 3) Rossi, P.E. et. al: Bayesian Statistics and Marketing, Wiley, 2005.
- 4) 国土交通省：『第4回全国幹線旅客純流動調査』，<http://www.mlit.go.jp/seisakutokatsu/jyunryuudou/index.html>, 2005.
- 5) (株)日経リサーチ：『2008地域ブランド戦略サーベイ（地域総合評価編）』，(株)日経リサーチ，2008.
- 6) Nakanishi, M. and Cooper, L.G. : Parameter Estimation for a Multiplicative Interaction Model - Least Square Approach, Journal of Marketing Research, XI, 303-311, 1974.