

# DEAの分析結果の統計による検討 (1)

## — 47都道府県の滞在型観光の分析 —

新 村 秀 一

### 1. はじめに

Data Envelopment Analysis (DEA, 包絡分析) 法は、「企業の経営効率性分析」とも呼ばれているが、評価を比率尺度で捕らえて可視化できる手法と考えれば良いであろう。DEAは線形計画法 (Linear Programming, LP) を評価対象 (Decision Making Unit, DMU) が $n$ 個あれば $n$ 回連続的に解くことで最適解が得られる。数理計画法は関数の最大/最小値を求める最適化手法 [2] [7] のため、統計とは無縁と考えられてきた。しかし筆者が行ってきた判別分析の研究 [3] [4] [5] では大きな貢献をした。また院生の岡野が統計の授業のレポートとして行った「普通車と小型車の2群判別」データは、DEAで日本車44車種の価格性能比の分析が行える [12]。DEAは重回帰型のデータの新しい手法 [9] と捕らえて、統計手法と共生するのが今後必要と考える。

数理計画法の重要な応用例を紹介した解説書 [8] の5章で、DEA研究の出発点であるCCRモデルを紹介した。そして数理計画法ソルバーのLINGO [2] [7] で作成した汎用モデルでExcel上のセル範囲Fから入出力データ $F(n*(p+q))$ を入力し実行するだけで、最適な重み $W(n*(p+q))$ とDEA効率値SCORE $(n*1)$ とクロス効率値 $S(n*n)$ がExcel上のセル範囲W, SCORE, Sに出力できる。従来の多くのDEA研究ではDEA効率値だけに注目がいき、重みWやクロス効率値Sの利用が考えられていない。重みの持つ情報を利用するためには、各入出力変数の最大値を1以上10未満に単位変換することが重要である。最大値で割って正規化すると単位がなくなり解釈に困るからである。またCCRモデルの利用法として、2入力1出力あるいは1入力2出力モデルで、2個の1入力1出力の比を考えて、その散布図でCCRモデルのDEA効率フロンティアを紹介し、各DMUは1象限でそのフロンティアの原点側に包み込まれると紹介している。そして非効率なDMUと原点を結ぶ直線とDEA効率フロンティアの交点が、そのDMUの改善目標と説明している。この説明は現実的でない。N個あるDMUの中で非効率なものも交点を求める手間だけでなく、そのような実体のない理想点は現実問題で改善目標になりえない。そこで、クロス効率値からDEAクラスターと呼ぶ考え方で改善目標になるDMUのパターンに分けることを提案した [9]。しかしDEA効率的なDMUが $k$ 個あれば、最大 $(2^k - 1)$ 個のDEAクラスターが生じる。改善活動を、最初からそのような小集団に分けて考えることは効果的とは考えられない。そこで、杉山 [14] の提案したInverted DEAモデ

ルの利用を考えた。CCRモデルでDEA効率的で、Inverted DEAでDEA非効率値が最大になるものをまず全体の改善目標とする考え方である。そして入力変数の1個を固定し、この改善目標の構成比率になるように他の変数値を修正すると、改善目標と同じ構成比率になる。全てのDMUがこの値を達成すれば、改善目標と同じ基準でDEA効率的になる。そして元の値からこの値を引いた差が、具体的な改善量と考えることができる。以上の[8]で紹介した方法に付け加えた新しい改善点は、「東京都23区の区立図書館のデータ」を用いて解説した[11]でまとめて紹介してあるので、[8]を踏まえて[11]の方法で今後DEAの分析を行えばよいと考える。DEA研究では、種々の精緻なモデルが開発されているが、現実応用ではDEA研究の発端となったCCRとInverted DEAモデルを十分使いこなすことである。

前著[13]の「日本の金融機関20行の分析」では、その上でCCRモデルとInverted DEAモデルで得られたDEA効率的とDEA非効率的な情報のSCOREとSCORE2を元データに反映させ、元データの変数と $p \times q$ 個の比率データとSCOREとSCORE2を一元配置の分散分析を行い、主成分分析で各DMUの布置の関係を分析し、最後にDEA効率値と非効率値のSCOREとSCORE2を元の変数と比率で重回帰することを提案した。これは、SCOREとSCORE2が元データで説明しがたく、比率情報でより説明しやすいかを複数のデータで調べる目的である。

本研究では、47都道府県の滞在型の観光を取り上げる。しかしこれまで扱った組織の分析では、従業員数、床面積、蔵書数などの入力変数は「1入力固定改善法」で示される改善量が改善策に利用できるのに対し、県民所得や宿泊施設の部屋数などは制御しにくい特徴があり、定性的な分析結果になることを区立図書館の分析結果と対比し注意する必要がある。その上で、統計手法の分析結果で単純なDEAによる解釈の間違いを訂正することが重要なことを示す。

統計分析にはJMP[1][6][10]を用いた。前期開講の「経営科学」では就活のため出席点の不足している4年次生には、就職を考えた企業10社程度と3入力3出力程度(資本金、従業員数、借入金、経常利益、純利益、売り上げなど)の分析を行い、5頁以上のDEA分析だけのレポートを課した。製粉企業に就職した学生は、圧倒的に日清製粉が良いことは分かっていたが、徳島製粉もDEA効率的になり、時間がなくて特定品目に特化しているであろうという結論を引き出した(実際には金ちゃんヌードルの販売)。メガバンクに就職した学生は、最初操作を間違っていてHPを引用し千葉銀行の取り組みがいかに優れているかを具体的な分析結果に基づかないレポートを提出した。正しい操作法を教えることで無事自分の就職先のメガバンクが千葉銀行より効率的なことを理解して安堵できた。統計分析の結果の解釈には、熟練した経験と知識が要る。しかしDEAはDEA効率的な企業とそうでない企業の白黒がはっきりするので、学生でも問題点を見つけやすく、その理由を考えさせることは就職先の知識を持ち合わせていない学生を社会に送り出すのに適切なテーマと考える。できれば、

日本中の大学で就活の一環として利用してもらいたいと考えている。

## 2. DEAによる47都道府県の滞在型観光業に関する分析を例として

平井[16]は、表1に示す47都道府県の一人当たり県民所得（100万円）、一人当たり他県民所得（100万円）、ホテル・旅館1施設あたり客室数、面積（km<sup>2</sup>）あたり温泉地数、面積（km<sup>2</sup>）あたり観光施設数を5個の入力変数と考え、県内観光客（100万人泊）、県外観光客（100万人泊）、外国人観光客（100万人泊）の3個の出力変数を用いて、DEAで47都道府県の効率性をCCRモデルだけで分析した先駆的研究と評価できる。Googleで検索しても現時点では観光に関するDEAによるアプローチはこの研究以外見当たらない。これは平井も指摘しているが、観光に関するデータの作成に手間隙がかかることが大きな点と考えられる。その点で、彼の公開したデータを利用して観光に関する分析を試みることは、この分野の研究の入門として意義があると考えられる。

表1 47都道府県の4入力3出力<sup>1</sup>

都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1 北海道	2.463	31.071	0.003	0.023	6.292	10.399	1.82
2 青森	2.443	21.091	0.016	0.043	0.386	1.078	0.03
3 岩手	2.346	21.563	0.006	0.031	1.025	1.728	0.09
4 宮城	2.615	31.487	0.008	0.092	1.783	2.105	0.08
5 秋田	2.334	24.836	0.011	0.046	0.479	1.343	0.03
6 山形	2.472	20.947	0.014	0.068	1.017	1.439	0.02
7 福島	2.775	19.617	0.010	0.051	1.715	3.590	0.08
8 茨城	2.843	19.337	0.007	0.086	0.433	0.890	0
9 栃木	3.104	20.274	0.011	0.111	1.007	3.960	0.08
10 群馬	2.921	23.337	0.016	0.106	1.001	4.068	0.03
11 埼玉	2.961	24.768	0.005	0.175	0.202	0.256	0.01
12 千葉	2.962	25.620	0.017	0.130	1.384	8.435	1.02
13 東京	4.820	63.007	0.011	0.393	1.379	3.879	2.29
14 神奈川	3.257	27.245	0.015	0.319	1.544	3.282	0.30
15 新潟	2.734	17.018	0.014	0.088	1.470	2.515	0.04
16 富山	3.013	27.411	0.035	0.166	0.634	0.973	0.06
17 石川	2.806	27.413	0.015	0.087	0.991	2.538	0.11
18 福井	2.819	13.879	0.010	0.087	0.472	0.904	0.02
19 山梨	2.773	16.045	0.007	0.137	0.368	2.834	0.32
20 長野	2.789	20.644	0.018	0.151	1.729	6.387	0.19
21 岐阜	2.863	17.887	0.007	0.084	0.504	2.155	0.13
22 静岡	3.389	18.351	0.016	0.129	1.619	7.525	0.27

<sup>1</sup> 一人当たり他県民所得の計算が分からないので分析から省く。県民所得は2006年度、その他は2007年度のデータで、観光庁、経済社会総合研究所、自然環境局および日本観光協会より平井が作成。ただし所得の単位表記は誤記と考え修正した。また温泉地数と観光施設数は、総数にすべきであるが、[16]との分析法の違いを比較するために修正はしない。

23	愛知	3.509	33.059	0.007	0.121	1.311	0.828	0.10
24	三重	3.193	17.042	0.010	0.093	0.652	2.773	0.05
25	滋賀	3.352	21.081	0.006	0.108	0.294	1.474	0.10
26	京都	2.976	30.569	0.005	0.111	0.718	6.403	0.87
27	大阪	3.083	52.551	0.016	0.170	0.716	2.780	0.96
28	兵庫	2.882	22.663	0.009	0.134	1.526	2.901	0.15
29	奈良	2.692	17.785	0.010	0.077	0.127	0.914	0.05
30	和歌山	2.665	18.774	0.010	0.089	0.389	2.346	0.10
31	鳥取	2.422	21.966	0.004	0.078	0.186	1.097	0.01
32	島根	2.437	19.686	0.007	0.057	0.212	0.908	0.01
33	岡山	2.800	22.967	0.006	0.072	0.399	1.026	0.02
34	広島	3.095	29.975	0.008	0.976	0.499	0.899	0.04
35	山口	2.883	19.364	0.010	0.087	0.536	1.216	0.01
36	徳島	2.694	14.234	0.008	0.054	0.062	0.433	0
37	香川	2.718	25.833	0.019	0.168	0.147	0.875	0.02
38	愛媛	2.487	26.623	0.008	0.081	0.183	0.870	0.02
39	高知	2.170	21.189	0.005	0.041	0.199	0.828	0.01
40	福岡	2.665	35.924	0.010	0.127	0.888	0.999	0.22
41	佐賀	2.475	22.716	0.011	0.118	0.296	0.973	0.03
42	長崎	2.159	26.408	0.009	0.104	0.453	2.361	0.41
43	熊本	2.398	17.608	0.011	0.077	0.652	2.371	0.37
44	大分	2.594	18.274	0.915	0.086	0.534	2.461	0.32
45	宮崎	2.150	29.399	0.008	0.048	0.212	0.862	0.08
46	鹿児島	2.283	21.983	0.011	0.072	0.556	1.785	0.08
47	沖縄	2.089	27.088	0.002	0.105	0.988	9.733	0.20
最大値		4.82	63.007	0.915	0.976	6.292	10.399	2.29

平井は表の4入力と、県内観光客と県外観光客をそれぞれ1出力とする4入力1出力の2個の分析と、所得を省いた3入力と外国人観光客を1出力とする1個の分析、そして他県民所得を含む5入力3出力の1個の計4個の分析を行い、DEA効率値のみの検討を行い全ての分析で北海道がDEA効率的であったと報告している。利用したDEA法ソフトの制約が大きく、データの持つ情報がCCRモデルで十分に分析されていないのは残念である。本研究では、新村が作成したDEA法の汎用ソフトと、1入力固定改善法という画期的な手法を用いた結果を示す。ただし入力変数の制御が困難であり、具体的な改善数値を達成することは困難であり、単に定性的な分析という歯切れが悪い結果になる。本研究は、今後DEAの分析例が増えた場合の反面教師と考えている。ただし、東京都の区立図書館で見たように、一見制御困難と考えられる床面積を半分減らすことは、その分を他目的に転用することで簡単に解決できる例もあることは重要である。そして、間違ったDEAの分析を避ける上でDEAで得られた豊富な情報を統計分析することを提案する。

### 3. DEA法に関する問題点の提起

入力を $x$ として出力 $y$ が成果を表す場合、単回帰式 ( $y = a*x+c$ ) で評価項目 $y$ が $x$ で予測で

きる。DEA研究者の多くが、回帰分析の欠点としてはn個の評価対象から共通の回帰係数aと定数項cを求めている点を指摘しているが適切ではない。重回帰分析に限らず統計手法はyの値が大きなものに注目する傾向があるが、DEAは入出力の値がともに小さくても比率で考えれば高いものを発見できることが重要である。例えば統計手法による評価では、企業における事業部評価でも、売り上げや利益規模の大きな中核事業が注目され、たとえ採算性が良くても規模の小さい事業部は評価されないことが多い点が問題である。これに対してDEAは、入力と出力の比を効率値 $y/x$ としてとらえることを提案した。その上で個々の $DMU_i$  (評価対象) に最適な重みを与えて、DEA効率値 $b_i*y_i/a_i*x_i$ を他の $DMU_j$  ( $j=1,\dots,n$ ) のクロス効率値<sup>2</sup>を1以下にするという制約のもとで最大化することを提案した。DEAの基本的なこの手法は、米国テキサス大学のCharnesとCooper両教授とRhodesによって開発されたのでCCRモデルと呼ばれている。下線を引いたものは、DEA効率値を求めたい $DMU_i$ に関係していることに注意してほしい。

$$\text{MAX} = \underline{b_i*y_i/a_i*x_i} ; \quad \underline{b_j*y_j/a_j*x_j} \leq 1 ; \quad j=1,\dots,n \quad \text{for } i \quad (1)$$

p個の入力とq個の出力がある場合、入出力と重みをベクトルに置き換えて、DEA効率値を $\underline{b_i*y_i/a_i*x_i}$ と定義すれば式(2)で一般化される。これによって重回帰分析で扱えない複数の出力変数も分析できることを統計手法に対する優位点としている。また企業などで比率尺度による評価が良く用いられている。しかしp入力q出力のデータがあれば、 $p*q$ 個の比率ができて、それぞれの比率で最大になるものが異なる可能性があり、総合的な判断が行えない。これを重みで総合化された入力と出力の比で評価できる利点に注目すべきであろう。

$$\text{MAX} = \underline{b_i*y_i/a_i*x_i} ; \quad \underline{b_j*y_j/a_j*x_j} \leq 1 ; \quad j=1,\dots,n \quad \text{for } i \quad (2)$$

このCCRモデルを用いる最大の利点は、評価の可視化と公平性が実現できる点と筆者は考える。すなわち評価対象と考えられる $DMU_i$ に最適な重みを求めているが、その結果DEA効率値が1になる場合と、ならない場合がある。従来の企業における評価法は、上司や専門家の経験や知識に負うところが大きい。そして、その基準が分かりにくく不明であり、評価が良くない場合には評価対象にとって与えられた評価が納得しにくかった。しかしDEAでは、評価対象自身に最適な重みを求めてなおかつ非効率であれば、その重みでクロス効率値が1になる他のDMUがいることになる。その場合、そのDMUを参照集合(手本)として改善点

<sup>2</sup> DEAの目的関数の値をDEA効率値、制約式で計算されるものをクロス効率値と区別する。

を考えることができる。これが重回帰分析のように共通の重みであったり、他のDMUの重みであったり評価基準があいまいであったりしない点が、評価の可視化や公平性を考える上で重要になる。

一方CCRモデルは各評価対象に一番有利な評価を行うため、入出力の変数が増えてくるとDEA効率的になる手本が増える問題がある。DEAの普及を考える場合、たくさん出てくる手本の中で一つの評価対象を手本にしてまず問題点(改善点)を検討し、必要であれば別の手本で追加検討する方が普及しやすい。またCCRモデルの欠点は、手本が全て1に抑えられて優先順位がつけられない点である。そこで式(3)のInverted DEAモデル[14]の利用を考えた。DEA効率値に代わってDEA非効率値のSCORE2 ( $\mathbf{b}_i^* \mathbf{y}_i / \mathbf{a}_i^* \mathbf{x}_i$ ) を考え、この重みを用いて47都道府県のDMU<sub>j</sub>が1以上になるという制約で最小化する重みを求める。このモデルの有用性は、DEA非効率値(SCORE2)が1になる非効率な手本に注目することではなく、CCRモデルで手本になった評価対象の中でDEA非効率値が最大の評価対象を最初の改善目標にできる点である。

$$\text{MIN} = \mathbf{b}_i^* \mathbf{y}_i / \mathbf{a}_i^* \mathbf{x}_i ; \quad \mathbf{b}_i^* \mathbf{y}_j / \mathbf{a}_i^* \mathbf{x}_j \geq 1 ; \quad j=1, \dots, n \quad (3)$$

さらにDEAを経営効率性の改善法として普及するために以下の点を提案している。

- ・DEAを普及するためには、最初の段階ではCCRとInverted DEAという基本モデルに限定したほうが良い。最新の研究成果までを考えると、多くの人の理解を得ることが難しく普及を困難にする。
- ・DEAは数理計画法で定式化され、多くのモデルが研究されている。しかし、普及のために数理計画法の理解を前提とせず、与えられた分析対象のデータF(Factor)と、最適化で得られた重みWとW2, そこから計算されたクロス効率値SとS2, DEA効率値SCOREと非効率値SCORE2といったデータで説明した方が、統計分析の知識がある多くの人の理解が得やすい。
- ・Excel上にデータFを与えるだけで、モデルのサイズに影響を受けないCCRとInverted DEAモデルが簡単に実行できる汎用モデルを開発した[8][11]。
- ・Fの各変数の最大値を1以上10未満になるように単位変換することで、数値計算上の問題の回避と重みの解釈が容易になる工夫をした。本研究では、最大値は63から0.9の範囲なので単位の修正は行わないことにした。
- ・1入力2出力あるいは2入力1出力モデルの場合、入力と出力の2個の比を作り散布図を描くことで、効率的DEAフロンティアとそれに包み込まれる非効率なDMUの改善目標がわかる。ただし入出力変数の和が4個以上になると3個以上の比があるので複数の散布図になり

解釈が困難になる。そこでクロス効率値から求めたDEAクラスターで対応することを提案した[9]。これは $p$ 入力と $q$ 出力から $p*q$ 個の比尺度ができて、最大 $p*q$ 個のDEAクラスターができて $n$ 個の評価対象を $p*q$ 個の小集団による改善活動を目指すことになる。しかし細分化された改善活動は現実的でなく、Inverted DEAモデルで改善目標を一つに決定することを考えた。またSCOREとSCORE2の値の組み合わせでDMUを4カテゴリーに分けた。

- ・DEAは、これまでの企業における評価で単に規模が小さいことで注目されなかった評価対象であっても、DEA効率値が1であれば手本であることを示してくれる。しかし、変数が多くなっていくと手本やDEAクラスターが増えていく傾向がある。評価対象全体でまず改善策を考える場合は、CCRモデルで効率的であり、Inverted DEAモデルで最大の非効率値(逆SCORE)をもつ評価対象を共通の改善目標と考えた方がよい。それがうまくいった後で、次の改善策を考えるべきである。
- ・改善方法を考える場合、評価対象のデータFを用いて、Excelで簡単に計算できる「1入力固定改善法」を提案する。これによって経済学部のある学生に就活希望の業種の企業を10社ほど集めさせて分析させたところ、データの入力に1時間、分析に2時間ほどで、社会経験のない学生でも容易に有益な知識が得られる。テーマの選定を自由にすると決められない学生が多いが、就職希望対象企業というように限定すれば、多くの学生のネックを解消できることが分かった。
- ・専門用語としての「DMU(意思決定主体)」と「参照集合」に代わって、柔らかい印象を与える「評価対象」と「手本」に置き換えて普及した方がよいと考える。
- ・そして、これらの出力結果を統計分析することで、より多くの人に理解してもらえることを示す。また1入力と1出力の比がもつ情報は元のデータにない情報を持っているような感触を得ている。 $p*q$ 個の比を統計分析することで、元データで得られない結果が得られるのではないかと考えている。

## 4. CCRモデルによる分析

### 4.1 DEA効率値と重みとクロス効率値

Excel上に表1の4入力3出力のデータを与え、セル範囲名をFとする。これを汎用LINGOモデルの「F=@OLE();」でLINGOに入力し配列FとしてLP計算に使い、CALC節で47回LP計算をする。そして、CCRモデルを実行しDEA効率値SCOREと重みW、クロス効率値Sを計算し、Excel上に指定したセル範囲名に「@OLE()=SCORE, W, S:」で出力する。その後で、Inverted DEAを連続実行する。

表2は、DEA効率値と重みである。W1からW7は表1の4入力3出力に対応した47都道府県の47個の重みである。1行目は北海道のDEA効率値を最大化する重みである。この重みを用

いた北海道のDEA効率値がSCORE欄に示される1である。この重みを北海道を含む47都道府県のデータに適用し1を越えないように制約する。2行目は青森の重みであり、青森のDEA効率値は0.15になる。この理由は青森に最適な重みで、他の県のクロス効率値が1になるためであり、青森はクロス効率値でこれらの県を特定して改善目標にする必要がある。47個のCCRモデルの分析で、北海道と静岡と沖縄の3県がDEA効率的であり参照集合すなわち47都道府県の手本になることが分かる。

しかしこの解釈は、現実応用において大きな誤解を生む。沖縄はあくまで沖縄を手本とするDEAクラスターの狭い範囲でDEA効率的である。Inverted DEAでは沖縄より例えば千葉のDEA非効率値が大きい、千葉はDEA効率値が1にはならない。これは北海道と比較して効率的と考えられないからである。即ちCCRモデルによるDEA効率値SCOREとInverted DEAモデルのSCORE2によるDEA非効率値の両方で判断する必要がある。この視点が従来のDEA研究では軽視されているように感じられる。

更に重みの解釈に関する指摘が少ない。CCRモデルでは入力値が小さく出力が大きければDEA効率値は大きくなる。このため4入力に対応するW1からW4で0になるのは相対的に経営資源をたくさん使い、0でない変数は使っていないことを示す。出力に対応するW5からW7は経営成果の小さな変数に0が与えられる。例えば北海道の重みは、県民所得、客室、温泉地の値が比較的大きいので重みは0になり、観光施設は少ないので43.5が与えられる。出力は県内客が多く、県外客や外国人客が比較的少ないので0になっている。ただしこの解釈にも注意が要る。表1の北海道と最大値の比は、7変数で2.463/4.82 : 31.071/63.007 : 0.003/0.915 : 0.023/0.976 : 6.292/6.292 : 10.399/10.399 : 1.82/2.29=0.51 : 0.49 : 0.003 : 0.024 : 1 : 1 : 0.79になる。前述の解釈に従えば、入力は観光施設ではなく温泉地のほうが小さく、出力では県内客に加えて県外客も重みが0以外になってよさそうであるが、総合化された入力値と出力値のため1変数単独で解釈できない。一方この分析結果と一般的な情報をつき合わせて検討すると、確かに北海道を訪れる外国人観光客は、自然や夕張メロンの食べ放題を目的としているというニュースが良く紹介されている。そこには旧来の観光施設や温泉地の影は薄い。また県内観光客が多いという点では、面積では他の県の4-5県分に相当するので北海道の県内観光客は農産物出荷量と同じく異常値になるので当たり前と考えられる。このようにDEAの分析結果は白黒がはっきりし、学生にも問題点が明確になる。その上で、問題点が正しいか、実際に検討させることが重要である。データに基づく客観的な数値から事実に基づいた結論を模索させることが重要である。

沖縄の重みから、温泉地が極端に少なく、県外客が多い特徴がある。これは調査分析するまでもなく正しいであろう。面積が小さいので滞在型の県内観光客の需要は小さいという他府県と異なった特徴がある。学生は全ての重みを分析する必要はなく、分かりやすいものだ



けを行えばよい。観光にかかわる職業については、詳細を分析すればよい。静岡の重みから客室数が少なく観光施設数が少ないが県外客が多いことになり、世界遺産富士山の観光客が突出しているか調べる必要がある。山梨は、客室が少ないのに県内と外国人観光客が多いことを示している。1施設あたりの客室数でなく山梨県の総客室数は少ないと考えられるので、総客室数を用いれば結果も影響を受けそうだ。また、外国人観光客では静岡に勝っているのは、忍野八海や湖が多いことが影響しているのか調べる必要がある。

Inverted DEAの重みW2は、入出力の比を最小化するので、入力重みが0になるのは使用する経営資源が少ないもの、出力の経営成果は多い変数に0が与えられる。しかし非効率なDMUに注目するのではなく値が大きなものに注目するので解釈には注意が要る。

表2 DEA効率値と重み

都府県	都道府県	SCORE	W1	W2	W3	W4	W5	W6	W7
1	北海道	1	0	0	0	43.5	0.16	0	0
2	青森	0.15	0	0.04	0	1.62	0	0.14	0
3	岩手	0.24	0	0.05	0	0	0.07	0.1	0
4	宮城	0.28	0	0.03	0	0	0.16	0	0
5	秋田	0.16	0.03	0.04	0	0.93	0	0.12	0
6	山形	0.24	0	0.05	0	0	0.24	0	0
7	福島	0.51	0	0.05	0	0	0.07	0.11	0
8	茨城	0.13	0	0.05	0	0	0.07	0.11	0
9	栃木	0.51	0	0.04	9.11	0	0.02	0.12	0
10	群馬	0.46	0.03	0.04	0	0.91	0	0.11	0
11	埼玉	0.04	0	0.04	0	0	0.2	0	0
12	千葉	0.91	0	0.04	0	0	0	0.09	0.15
13	東京	0.64	0.21	0	0	0	0	0	0.28
14	神奈川	0.34	0	0.04	0	0	0.05	0.08	0
15	新潟	0.44	0	0.06	0	0	0.08	0.13	0
16	富山	0.11	0	0.04	0	0	0.18	0	0
17	石川	0.26	0.05	0.03	0	0	0.01	0.1	0
18	福井	0.19	0	0.07	0	0	0.1	0.15	0
19	山梨	0.48	0	0.06	0	0	0	0.14	0.25
20	長野	0.81	0.07	0.04	0	0	0.02	0.12	0
21	岐阜	0.32	0	0.05	9.92	0	0	0.15	0.07
22	静岡	1	0	0.04	0	1.59	0	0.13	0
23	愛知	0.2	0	0.03	0	0	0.15	0	0
24	三重	0.42	0	0.05	4.78	1.22	0	0.15	0
25	滋賀	0.19	0	0.04	8.58	0	0	0.13	0.06
26	京都	0.61	0	0.03	6.05	0	0	0.09	0.05
27	大阪	0.42	0.32	0	0	0	0	0	0.44
28	兵庫	0.37	0	0.04	0	0	0.06	0.09	0
29	奈良	0.14	0	0.05	0	1.78	0	0.15	0
30	和歌山	0.33	0	0.05	4.43	1.13	0	0.14	0
31	鳥取	0.14	0	0.04	4	1.02	0	0.13	0
32	島根	0.13	0	0.05	4.46	1.14	0	0.14	0

33	岡山	0.13	0	0.04	8.49	0	0.02	0.11	0
34	広島	0.09	0	0.03	0	0	0.05	0.07	0
35	山口	0.17	0	0.05	0	0	0.07	0.11	0
36	徳島	0.08	0	0.06	0	2.27	0	0.19	0
37	香川	0.09	0.06	0.03	0	0	0.02	0.1	0
38	愛媛	0.09	0.02	0.03	0	0.84	0	0.11	0
39	高知	0.11	0	0.04	4.28	1.09	0	0.14	0
40	福岡	0.13	0.38	0	0	0	0.15	0	0
41	佐賀	0.12	0.06	0.04	0	0	0.02	0.12	0
42	長崎	0.27	0	0.04	0	0	0	0.09	0.15
43	熊本	0.39	0	0.06	0	0	0	0.13	0.23
44	大分	0.38	0	0.05	0	0	0	0.13	0.22
45	宮崎	0.09	0.44	0	0	1.14	0	0.11	0
46	鹿児島	0.23	0.03	0.04	0	1.01	0	0.13	0
47	沖縄	1	0	0	500	0	0	0.1	0

47県の重みから、47行\*47列のクロス効率値Sが計算される。この1列から47列に北海道から沖縄の重みを47県に適用したクロス効率値が出力される。表3はDEA効率的になる3県のみクロス効率値を表示したものである。S1はクロス効率値の1列目すなわち北海道の重みで47都道府県の効率値を計算し、3行だけを表示した。北海道だけが1であり、他の46都道府県は1未満である。他の46都道府県はある程度の成果があっても北海道に押さえられて効率的にならない。S2は青森の重みで計算したものであり、青森は0.15と低いのは、北海道と静岡がその重みでクロス効率値が1になるからである。青森は、北海道と静岡を手本にして改善することになる。この点はDEAクラスターの結果とあわせて考えればよい。北海道は46県の、静岡は39県の、沖縄は12県の手本になっていることが分かる。即ちDEA研究でDEA効率的であり参照集合即ち手本とする3県の間に、明らかに序列があることが十分に議論されていない。

表3 DEA効率的になる3県のみクロス効率値

都道府県	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
北海道	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
静岡	0.05	1	1	0.436	1	0.44	1	1	1	1	0.44	1	0.11	1	1	0.44
沖縄	0.03	0.97	0.82	0.18	1	0.18	0.82	0.82	1	1	0.18	0.85	0.13	0.82	0.82	0.18
都道府県	S17	S18	S19	S20	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31	S32
北海道	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
静岡	1	1	1	1	1	1	0.44	1	1	1	0.11	1	1	1	1	1
沖縄	1	0.82	0.85	1	1	0.97	0.18	1	1	1	0.13	0.82	0.97	1	1	1
都道府県	S33	S34	S35	S36	S37	S38	S39	S40	S41	S42	S43	S44	S45	S46	S47	
北海道	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.71	
静岡	1	1	1	1	1	1	1	0.19	1	1	1	1	0.49	1	0.1	
沖縄	1	0.82	0.82	0.97	1	1	1	0.19	1	0.85	0.85	0.85	1	1	1	

表4は、表3のクロス効率値から求めた筆者が開発したDEAクラスター[9]である。北海道だけが手本になる県は北海道を含む9県である。北海道と静岡を手本とするのは静岡を含む18県で、3県を手本とするのも18県ある。北海道と沖縄を手本とするのは宮崎だけで、沖縄を手本とするのは沖縄だけである。DEA研究では、これらの5つのグループで改善を考えることを提案している。しかし改善活動をこのように細分化して行うことは現実的でない。また4入力3出力モデルは、12個の1入力1出力モデルの比で表される空間で表現されることができると考えられる。この場合、(北海道, 静岡, 沖縄)のDEAクラスターは、北海道, 静岡, 沖縄の三角形がDEAフロンティアになり原点とできる三角錐に18県が入るか否かの理論的証明は今後の検討課題である。そして(北海道, 静岡)のDEAクラスターは、この2県と原点を結ぶ三角形上に18県が布置されるのか検討する必要がある。誰かこの問題を解決してもらえればと希望している。

表4 DEAクラスター

手本	数	構成県
北海道 (1)	9	1, 4, 6, 11, 13, 16, 23, 27, 40
北海道, 静岡 (22)	18	2, 3, 7, 8, 12, 14, 15, 18, 19, 22, 28, 29, 34, 35, 36, 42, 43, 44
北海道, 静岡, 沖縄	18	5, 9, 10, 17, 20, 21, 24-26, 30-33, 37-39, 41, 46
北海道, 沖縄	1	45
沖縄 (47)	1	47

## 5. Inverted DEAと1入力固定改善法

### 5.1 Inverted DEA

表5はInverted DEAモデルで計算したDEA非効率値(SCORE2)と重み(W<sub>2</sub>)である。クロス効率値S<sub>2</sub>の分析は省く。重みの後にCCRモデルのSCOREを付加し、SCOREとSCORE2の重みをカテゴリー化してGroup欄で4つのグループに分ける。AはDEA効率値が1の3県で、DEA非効率値は北海道が26.3と最大で富士山をもつ静岡が13.6で沖縄が8.3である。このようにCCRモデルが1でDEA効率的であったとしても、DEA非効率値で大きな違いが分かる。本研究では「1入力固定改善報」で取りあえずは北海道を手本に改善策を検討する。しかし統計分析で指摘するとおり、北海道よりも静岡を目標とした分析も行う。

DEA研究の趨勢は、数理計画法は種々のモデルを簡単に開発できるので、新しい手法モデルの開発に力点があるが、このように入門手法を組み合わせるだけで大きな成果が得られることを重要視していない問題がある。分類Dは、DEA非効率値が1の茨城、埼玉、広島(厳島神社)、徳島(阿波踊り)、香川(金比羅宮)、大分(温泉大国)の6県である。厳島神社を持つ広島、金比羅宮を持つ香川、温泉の多い大分が含まれているが、これらの観光資源が生かされていない県と考えられる。これに対してDEA非効率値が沖縄の8.3以上と大きい県は、

北海道 (26.8), 静岡 (13.6), 長野 (12.3), 福島 (11.1), 千葉 (10.8), 兵庫 (9), 栃木 (8.3), 沖縄が8.3である。すなわち沖縄はDEA効率的であるが分類Bの長野 (善光寺), 福島 (会津, 喜多方), 千葉 (デズニールランド), 兵庫 (港神戸と城之崎), 栃木 (日光) の5県が沖縄より大きい。これはこれらの5県がCCRモデルで北海道か静岡と比較してDEA効率値が押さえられているからである。以上から北海道を手本とした場合と, 静岡を手本とした「1入力固定改善法」の検討を行う。以上から分類のAはSCOREが1の3県, Bはそれ以外のSCORE2が沖縄の8.3以上の5県, Cは8.3未満0以上の33県, DはSCORE2 = 1の6県である。

表5 Inverted DEAモデルのDEA非効率値 (SCORE2) と重み (W\_2)

都道府県	SCORE2	W1	W2	W3	W4	W5	W6	W7	SCORE	GROUP
1 北海道	26.3	0	0.03	0	0	3.18	0.60	0	1	A
2 青森	2.7	0	0.04	4.11	0	4.39	0.90	0	0.15	C
3 岩手	6.2	0	0.05	4.19	0	4.48	0.92	0	0.24	C
4 宮城	6.4	0	0.03	7.57	0	0	3.03	0	0.28	C
5 秋田	2.9	0	0.04	3.59	0	3.83	0.79	0	0.16	C
6 山形	3.1	0	0.03	25.5	0	0	1.18	65.1	0.24	C
7 福島	11.1	0.28	0	22.0	0	0	1.92	50.0	0.51	B
8 茨城	1.0	0	0	0	11.5	0	0.24	263	0.13	C
9 栃木	8.3	0	0.03	4.19	2.35	5.42	0.72	0	0.51	B
10 群馬	5.5	0	0	33.7	4.34	0	0.74	91.4	0.46	C
11 埼玉	1.0	0	0.04	0	0	0	3.91	0	0.04	D
12 千葉	10.8	0	0.03	3.35	1.87	4.33	0.57	0	0.91	B
13 東京	3.2	0	0.01	0	0.75	1.76	0.21	0	0.64	C
14 神奈川	6.2	0	0.02	2.40	1.35	3.11	0.41	0	0.34	C
15 新潟	6.3	0	0.01	36.0	4.58	0	0.97	96.5	0.44	C
16 富山	2.7	0	0.03	3.02	0	3.23	0.67	0	0.11	C
17 石川	5.2	0	0.03	3.22	0	3.44	0.71	0	0.26	C
18 福井	2.7	0.28	0	21.7	0	0	1.90	49.4	0.19	C
19 山梨	3.6	0	0.03	0	4.17	9.76	0	0	0.48	C
20 長野	12.3	0	0.03	3.69	2.07	4.77	0.63	0	0.81	B
21 岐阜	5.0	0	0.06	0	0	5.53	1.05	0	0.32	C
22 静岡	13.6	0.20	0	4.12	2.08	3.43	1.07	0	1	A
23 愛知	2.4	0	0.03	7.28	0	0	2.92	0	0.2	C
24 三重	6.0	0.30	0	4.87	0	2.47	1.59	0	0.42	C
25 滋賀	2.6	0	0.04	0	2.36	5.54	0.66	0	0.19	C
26 京都	5.4	0	0.03	0	0	7.51	0	0	0.61	C
27 大阪	2.3	0	0.02	0	0	1.88	0.36	0	0.42	C
28 兵庫	9.0	0	0.04	3.95	0	4.22	0.87	0	0.37	B
29 奈良	1.6	0	0.03	6.06	5.3	12.3	0	0	0.14	C
30 和歌山	4.2	0	0.04	0	2.7	6.34	0.76	0	0.33	C
31 鳥取	1.8	0	0.05	0	0	4.50	0.85	0	0.14	C
32 島根	1.7	0	0.05	0	0	0	1.18	106.0	0.13	C
33 岡山	2.6	0	0.04	0	0	4.31	0.81	0	0.13	C
34 広島	1.0	0	0.01	0	0.86	0	1.11	0	0.09	D
35 山口	2.6	0	0.04	30	0	0	1.39	76.7	0.17	C

36	徳島	1.0	0	0.07	0	0	16.10	0	0	0.08	D
37	香川	1.0	0	0.02	3.41	2.98	6.90	0	0	0.09	D
38	愛媛	1.3	0	0.04	0	0	3.71	0.70	0	0.09	C
39	高知	1.7	0	0.05	0	0	4.67	0.88	0	0.11	C
40	福岡	2.6	0	0.03	6.59	0	0	2.64	0	0.13	C
41	佐賀	2.1	0	0.04	3.91	0	4.17	0.86	0	0.12	C
42	長崎	3.4	0	0.04	0	0	3.75	0.71	0	0.27	C
43	熊本	6.1	0	0.05	4.98	0	5.32	1.10	0	0.39	C
44	大分	1.0	0	0	1.09	0	0	0.01	3.06	0.38	D
45	宮崎	1.3	0	0.03	0	0	3.36	0.64	0	0.09	C
46	鹿児島	4.0	0	0.04	4.03	0	4.31	0.89	0	0.23	C
47	沖縄	8.3	0	0.02	0	3.59	8.41	0	0	1	A

## 5.2 1入力固定改善法（北海道を手本とした場合）

北海道を手本とする場合、北海道の構成比率を改善目標と考える。そして47都道府県の4個の入力の1個を固定し、残りの6変数の値を、北海道と同じ構成比率になるように変更すればDEA効率値は全て1になる。現在の値からこの改善目標を引いたものでもって改善すべき問題点と改善する数値目標が分かる。

表6は県民所得を固定した場合で、北海道の行と県民所得の列は0になる。沖縄の県外観光客だけが正であり、県外観光客に頼っていることが分かる。残りの45県は全て北海道と比較し県民所得あたりの観光客が全て少ない。北海道は他の都府県より面積が広く県内観光客が多くなる特徴がある。そして、スキーを含む冬の観光資源で県外と外国人観光客を誘致できる利点がある。入力のカ室数が正のものは人口の多い、東京、大阪、福岡のほか、宮崎と沖縄が正である。前者はビジネス利用が含まれ、後者の2県は観光用であろう。これに対して温泉地と観光施設の多くは正であり、北海道に対して他県はこれらの比率が高いことになり、北海道は必ずしも $\text{km}^2$ あたりの温泉や観光施設に恵まれているわけではないことを示す。

表6 県民所得を固定した1入力固定改善法

都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1 北海道	0	0	0	0	0	0	0
2 青森	0	-9.7	0.01	0.02	-5.85	-9.2366	-1.8
13 東京	0	2.2	0.01	0.35	-10.9	-16.471	-1.3
27 大阪	0	13.7	0.01	0.14	-7.16	-10.237	-1.3
40 福岡	0	2.3	0.01	0.1	-5.92	-10.253	-1.7
45 宮崎	0	2.28	0.01	0.03	-5.28	-8.2155	-1.5
47 沖縄	0	0.74	-0	0.09	-4.35	0.91306	-1.3

表7は客室を固定した1入力固定改善法である。基本的には千葉県を含む40県が北海道に比べて客室比で県民所得、 $\text{km}^2$ あたりの温泉地と観光施設が多いのに観光客全てで劣っている。例えば距離と自然環境の似た青森は、北海道に比べ温泉地が多く、恐山や弘前城の桜、

三内丸山古墳と観光施設が多いのに、北海道に比べて観光客に劣っている。その理由としては、ビジネス客も含めた千歳空港の役割があろう。一方、集客力の多いデズニーランドを持つ千葉県は近郊各県からの日帰り客が多いという点も考えられる。これは京都や奈良といった観光地も近畿各県からの日帰り客の割合が多いことが考えられる。すなわち今回のデータは、日帰り観光の分析はできない。出力では沖縄だけが県外観光客が北海道基準より多い。入力の県民所得は県民所得の高い東京、大阪、福岡と宮崎、沖縄が負なのは客室数が他県より多いためであろう。温泉地と観光施設との比で考えて、多くの県が北海道より多いが、交通の便と北海道の北のイメージに負けていることが分かる。この点で、観光施設の多い九州が魅力をアピールすることが考えられる。

表7 客室を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	0	0	0	0	0	0	0
12	千葉	0.93	0	0.01	0.11	-3.8	-0.1396	-0.5
22	静岡	1.93	0	0.01	0.12	-2.1	1.38319	-0.8
13	東京	-0.2	0	0	0.35	-11.4	-17.209	-1.4
27	大阪	-1.1	0	0.01	0.13	-9.93	-14.808	-2.1
40	福岡	-0.2	0	0.01	0.1	-6.39	-11.024	-1.9
45	宮崎	-0.2	0	0.01	0.03	-5.74	-8.9774	-1.6
47	沖縄	-0.1	0	-0	0.08	-4.5	0.66705	-1.4

表8は温泉地を固定した1入力固定改善法である。沖縄だけが他の府県と異なり県民所得、客室、観光施設、県外観光客が少ない温泉に対し北海道よりも多く、その他の変数は負である。それ以外の45県は、青森のように温泉が多いため観光施設も負と全てが負になる15県（青森、岩手、秋田、山形、福島、群馬、千葉、新潟、富山、石川、徳島、熊本、大分、宮崎、鹿児島）と温泉の少ない30県に分かれる。

表8 温泉地を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	0	0	0	0	0	0	0
13	東京	-4.2	-50.9	0	0.31	-21.7	-34.251	-4.4
44	大分	-749	-9458	0	-6.9	-1919	-3169.2	-554
47	沖縄	0.45	6.374	0	0.09	-3.21	2.80033	-1

表9は観光地を固定した1入力固定改善法の結果である。大分が観光施設に対して温泉施設が日本一で、青森、秋田、山形、富山も多い。広島だけが温泉地が少ない。残り40県の温泉地は少数点3桁で四捨五入して0である。このことは、近年温泉地が観光客に対して影響を低下していることを示す。

表9 観光地を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	0	0	0	0	0	0	0
2	青森	-2.2	-37	0.01	0	-11.4	-18.364	-3.4
5	秋田	-2.6	-37.3	0.01	0	-12.1	-19.455	-3.6
6	山形	-4.8	-70.9	0.01	0	-17.6	-29.306	-5.3
16	富山	-15	-197	0.01	0	-44.8	-74.081	-13
34	広島	-101	-1289	-0.1	0	-267	-440.38	-77
44	大分	-6.6	-97.9	0.9	0	-23	-36.422	-6.5

### 5.3 1入力固定改善法(静岡を手本とした場合)

日本における滞在型観光ビジネスにおいて宿泊客による分析は北海道が突出しているので参考にならない。そこで静岡を手本にして分析を行う。比較するのは富士山を観光資源として共有する山梨、そして富士山と並んで日本三霊山を有する富山と石川、参考に北海道と沖縄と比較する。

表10は県民所得を固定した1入力固定改善法である。山梨は静岡と比較して客室が1.03室多く $\text{km}^2$ あたり観光施設も僅かであるが0.031箇所多く温泉地は0.006箇所少ないが、県内宿泊客は96万人泊、県外は369万人泊、外国人は10万人少ない。新幹線と私鉄のハンディキャップはあるが、静岡との差別化がうまく伝わっていないと考えられる。

一方、富山と石川は客室も11室以上、 $\text{km}^2$ あたり温泉地も0.002箇所以上、観光施設は富山が0.05箇所多いが、宿泊客は少ない。これは世界遺産である富士山と、三霊山という今日的な影響力の少なさ、交通の問題が考えられる。山梨以上に企画力があるが、石川は加賀文化を情報発信してきているようだ。北海道は温泉地や観光施設が少ないのに、宿泊客が多く客室もそれに答えているようだ。沖縄は、県内宿泊客が1万人泊少ないが、県外宿泊客が509万人はく、外国人が4万人泊多い。

表10 県民所得を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	0	17.73	-0.009	-0.071	5.115	4.93011	1.62
16	富山	0	11.1	0.021	0.051	-0.81	-5.717	¥1
17	石川	0	12.22	0.002	-0.020	-0.35	-3.6925	-0.1
19	山梨	0	1.03	-0.006	0.031	-0.96	-3.3232	0.1
22	静岡	0	0	0.000	0.000	0	0	0
47	沖縄	0	15.78	-0.008	0.025	-0.01	5.09454	0.04

表11は客室を固定した1入力固定改善法である。県民所得は46県全てが負で、温泉地は富山だけが0.0011と正で、観光施設は0.024箇所多いが、県内観光客は北海道が355万人泊、外

国人は北海道が135万人泊、山梨が8万人泊多いだけである。北海道は入力が全て負であるのに、宿泊客は静岡と比較しても圧倒的に多い。

表11 客室を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	-3.3	0	-0.024	-0.195	3.551	-2.342	1.35
16	富山	-2	0	0.011	-0.027	-1.78	-10.267	-0.3
17	石川	-2.3	0	-0.009	-0.106	-1.43	-8.703	-0.3
19	山梨	-0.2	0	-0.007	0.024	-1.05	-3.7454	0.08
22	静岡	0	0	0.000	0.000	0	0	0
47	沖縄	-2.9	0	-0.022	-0.085	-1.4	-1.3747	-0.2

表12は温泉地を固定した1入力固定改善法である。温泉地の少ない北海道、山梨、沖縄は静岡に比べて入力の県民所得と客室が相対的に正の値をとる。

表12 温泉地を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	1.83	27.63	0.000	-0.001	5.988	8.98806	1.76
16	富山	-4.4	-12.7	0.000	-0.116	-2.91	-15.488	-0.5
17	石川	-0.4	10.21	0.000	-0.034	-0.53	-4.5167	-0.1
19	山梨	1.29	8.016	0.000	0.081	-0.34	-0.4582	0.2
22	静岡	0	0	0.000	0.000	0	0	0
47	沖縄	1.67	24.79	0.000	0.089	0.786	8.79238	0.17

表13は観光施設を固定した1入力固定改善法である。山梨の客室と県外宿泊客が少ないのは別荘と保養所を統計データに含めているかいないか確認する必要がある。

表13 観光施設を固定した1入力固定改善法

都府県	都道府県	県民所得	客室	温泉地	観光施設	県内	県外	外国人
1	北海道	1.86	27.8	0.000	0.000	6.003	9.05733	1.77
16	富山	-1.3	3.797	0.014	0.000	-1.45	-8.7103	-0.3
17	石川	0.52	15.04	0.004	0.000	-0.1	-2.537	-0.1
19	山梨	-0.8	-3.44	-0.010	0.000	-1.35	-5.1577	0.03
22	静岡	0	0	0.000	0.000	0	0	0
47	沖縄	-0.7	12.15	-0.011	0.000	-0.33	3.608	-0

## 6. 統計分析による検討

### 6.1 GROUPによる一元配置

図1はSCOREとSCORE2の値で47都道府県を4つのグループに分け、これで各変数を1元



配置の分散分析することでDEAの分析結果を詳細に考察できる。DEAで用いる入力と出力の関係は少し癖がある。経営資源としての入力が多ければ、経営成果としての出力も多いという枠組み中であって、できるだけ少ない経営資源で多くの経営成果を達成できる効率的なDMUを探すのがDEAの目的と考えるべきであろう。あるいは同じ程度の経営資源の中で経営成果の大きなものを効率的と考えたほうが分かりやすい。

図1・1は、所得、客室、温泉地数である。所得に関しては、DEA分析で評価されないCの県の東京(13)について愛知(23)が高いことが分かる。これに対して北海道(1)は250万円と東京の半分であり、静岡(22)は340万円である。県民所得が多ければ、他府県や海外への観光が多くなるが、県内旅行ははっきりしない。この点を検討するため、県民所得のほぼ同じ北海道と東北6県を比較すると北海道の県内観光客が圧倒的に多いことをCCRモデルは表すが、北海道の面積は83457km<sup>2</sup>と東北6県の66889km<sup>2</sup>より大きい、東北6県の県内客の合計はほぼ北海道と同じである。[16]の欠点を解消するには、47都道府県の比較は問題があり、道州単位で行ったほうが良いかもしれないことが示唆される。

客室では大阪(27)がビジネスユースもあり東京についている。温泉地は大分(44)が圧倒的に大きいことが分かる。このような異常値は間違いがないか検討する必要がある。この分析結果で、DEAの結果を間違えて解釈していないか確認に用いることができる。しかし、滞在型の観光といえば温泉がイメージされるが、本分析ではかつての新婚旅行のメッカである宮崎は退潮傾向にあるようだ。

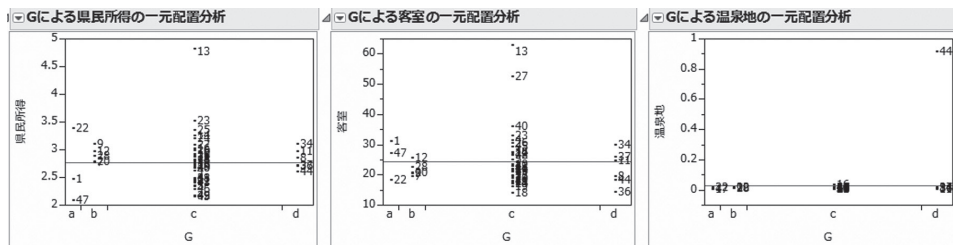


図1・1 県民所得、客室、温泉地

図1・2は観光施設、県内客、県外客である。観光施設数は、圧倒的にAとBが少なくCが一般的に大きな値である。Dは広島(34)を除けばCほど多くない。広島の観光施設数の多さは不思議である。また施設あたりの定義を見直す必要も考えられる。県内客は北海道だけが外れ値であり、これまでの主張を裏付けている。県外客は、 $A > B > C > D$ の順であり、これは常識に合致する。その意味で分析に用いていない他県民所得が、もしその県を観光にきた客の県民所得の重みつき平均値であれば重要な入力項目と考えられる。

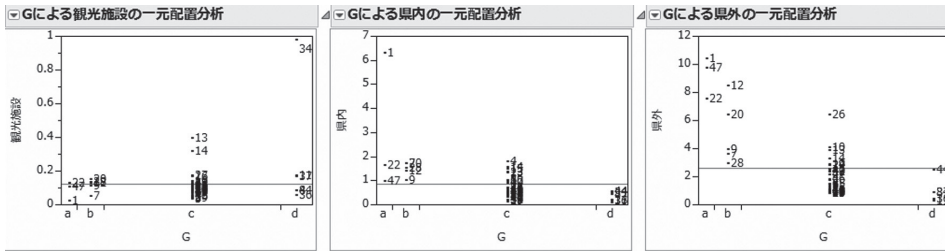


図1・2 観光施設, 県内容, 県外客

図1・3は、外国人客, Y1(県内容)/所得, Y1/客室の結果である。外国人客は東京(13)と千葉(12)が多いのは日本への玄関であり、大阪(27)は東南アジアの玄関でありそれが京都(26)へ流れていることが伺われる。これに対して北海道は確かにオーストラリアなどのスキー用の別荘や、夕張メロンや映画の舞台を訪れる東南アジアからの観光客の誘致に成功しているニュースからも分かる。しかし180万人を6県で割ると30万人になり、これを超えているのは千葉の102万人、東京の230万人、神奈川の30万人、山梨の32万人、静岡の27万人、京都の87万人、大阪の96万人、長崎の41万人、熊本の37万人、大分の32万人であり、外国人観光客の動向が分かる。しかし熊本をおとづれる外国人は阿蘇、大分は温泉が目的であるのか他に訪問地があるのか、調べる必要がある。最近では吉祥寺の「てんや」にも中国人ツアー客が見られるが、近くには本学しかないのので何を観光目的にしているか分からない。所得あたりの県内容は北海道だけが異常値である。客室あたりの県内容は、北海道のほか新潟(15)、長野(20)、静岡(22)の値が大きいが、都道府県の総客室数でなければ検討するのは保留すべきである。

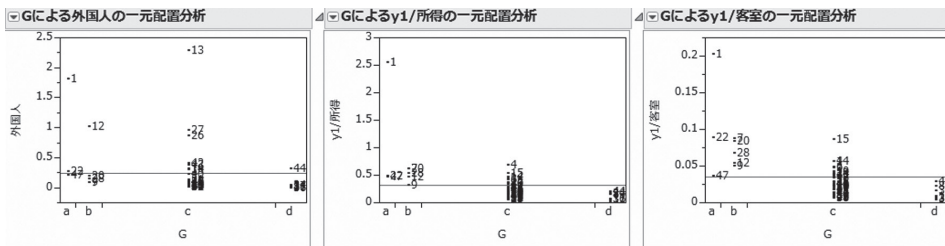


図1・3 外国人客, Y1/所得, Y1/客室

図1・4はY1/温泉, Y1/施設, Y2(県外客)/所得の結果である。温泉地あたりの県内容は、圧倒的に北海道と沖縄が大きい。km<sup>2</sup>あたりの温泉地の少ない北海道は大自然、沖縄はトロピカルな環境が、県内と県外客に大きな影響を持っているようだ。所得あたりの県外客は、北海道と沖縄に続いて、千葉(12)、長野(20)、静岡(22)、京都(26)が大きな値である。

ただしこの比率は、所得が高ければ県外客が多いという仮定は成り立たないであろう。

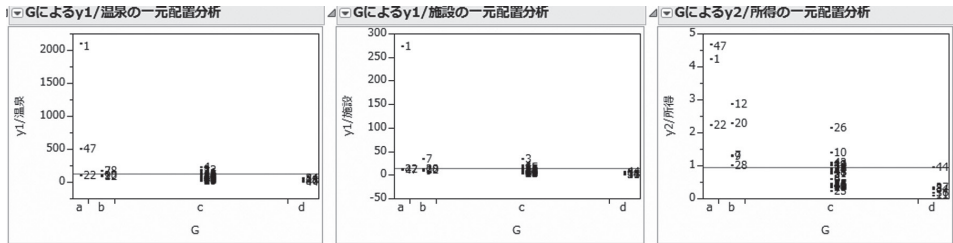


図1・4 Y1/温泉, Y1/施設, Y2/所得

図1・5はY2/客室, Y2/温泉, Y2/施設の結果である。温泉地あたりの県外客は沖縄と北海道が外れ値で、これは温泉地が観光資源になっていると解釈するのではなく、温泉地以外が北海道と沖縄の目玉になっていると考えるべきであろう。同様に施設あたりの県外客は北海道が外れ値で、旧来の観光用の施設以外が目玉と考えるべきであろう。

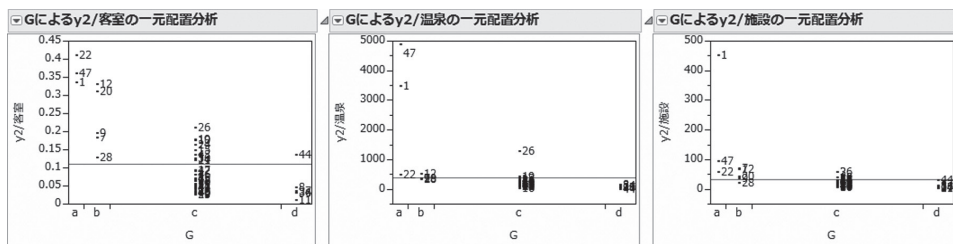


図1・5 Y2/客室, Y2/温泉, Y2/施設

図1・6はY3/所得, Y3/客室, Y3/温泉の結果である。所得と客室あたりの外国人客の解釈は無理に行わないほうが良いであろう。温泉地あたりの外国人客は、他と同じく温泉地でなく他の目玉があると考えべきであろう。

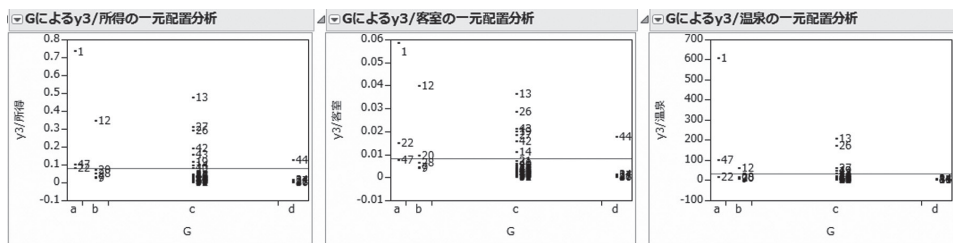


図1・6 Y3/所得, Y3/客室, Y3/温泉

図1・7はY3/施設の結果であり、温泉地と同じく施設が北海道の目玉になっていないことを表す。

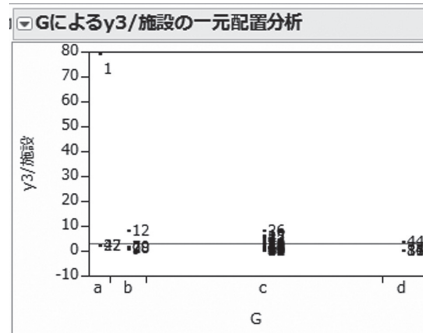


図1・7 Y3/施設

図1 一元配置の分散分析

## 6.2 主成分分析

図2は、JMP [1] [6]で表1の元データと SCOREと SCORE2を主成分分析し、左から固有値、第1主成分と第2主成分上でスコアプロットと因子負荷量を描いた。固有値が1以上は3個あり、9変数に対してデータのばらつきが大きいようだ。スコアプロットから、第1主成分の-2あたりに多くの県が集積し、第1象限の東京都(13)方向と第4象限の北海道(1)の方向に変数値が大きな値をとる都道府県が布置されているようだ。因子負荷量から第1象限は、東京都の値が大きい変数(県民所得、客室、観光施設、外国人)が布置している。第4象限には、県内と県外の宿泊客数とDEA効率値と非効率値が布置している。そして温泉地数は第3象限に布置し、近年宿泊を伴う観光の主流になっていないようだ。第4象限にはデズニーランドを抱える千葉(12)、静岡(22)、長野(20)、沖縄(47)、福島(7)が目につく。第1象限には、東京(13)に続いて大阪(27)、神奈川(14)、京都(26)といった関東と関西の中核府県が布置している。第2象限には、広島(34)が目につくがその理由は分からない。

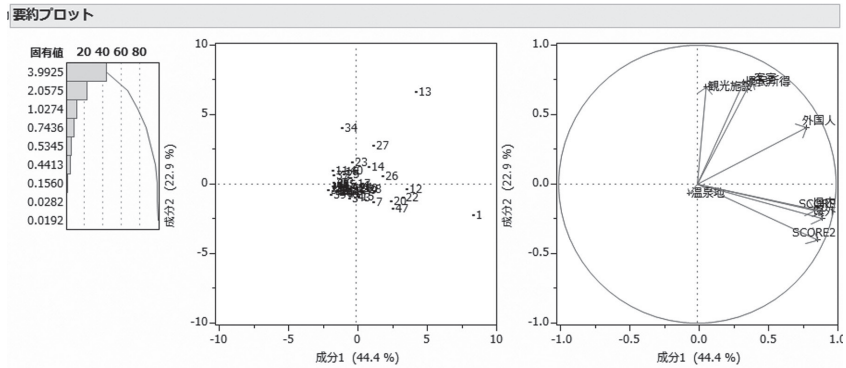


図2 元データとSCOREとSCORE2の主成分分析（左から固有値，スコアプロットと因子付加量）

表14は因子負荷量であり，因子負荷プロットで重なりがあって分かりづらいものの確認に用いればよい。分析に用いていないが第3主成分も示す。各象限で距離の近いものからクラスターすることが考えられる。

表14 元データの因子負荷量

行	主成分1	主成分2	主成分3	象限
県民所得	0.33	0.73	0.03	1
客室	0.39	0.75	-0.03	1
温泉地	-0.06	-0.07	0.99	3
観光施設	0.06	0.70	-0.02	1
県内	0.85	-0.19	-0.07	4
県外	0.90	-0.25	0.05	4
外国人	0.79	0.41	0.12	1
SCORE	0.90	-0.19	0.10	4
SCORE2	0.87	-0.40	-0.13	4

図3は，12個の比データと効率値の主成分分析による固有値，スコアプロットと因子負荷量である。固有値が1以上のものは12変数中2変数以上であり，最適化によってある程度ばらつきが2変数に凝縮されていることが分かる。スコアプロットから第3象限から第1象限に45度の方向に直線状に布置している。効率値という基準で最適化で変換され，沖縄(47)が目立っている。北海道は第4象限に外れ値として布置している。宿泊を伴う観光に関して，北海道は他の46県と異質である。これは農産物を都道府県で分析しても観測され，都道府県の農地面積で基準化するとある程度是正される。観光データでは何で基準化すればよいか検討中である。因子負荷プロットは分かりにくいので，表15で確認すると，効率値と第1象限に y1/客室，y2/所得，y2/客室，y2/温泉，y3/客室が布置されている。大きくは客室あたりの観

光客数が関係している。第4象限には、非効率値と y1/所得, y1/温泉地, y1/観光施設, y2/観光施設, y3/所得, y3/温泉地, y3/観光施設が布置している。大きくは所得, 観光施設, 温泉地あたりの観光客数が関係している。

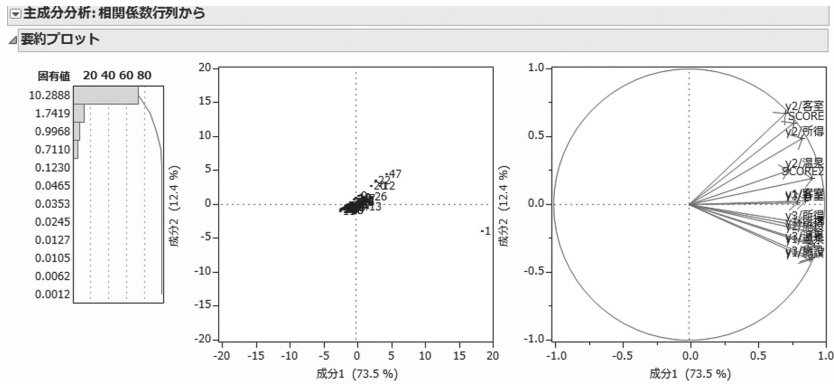


図3 非データの主成分分析 (左から固有値, スコアプロットと因子付加量)

表14は比データの因子付加量である。分析に用いていないが第3主成分も示す。各象限で距離の近いものからクラスターすることが考えられる。

表15 非データの因子付加量

行	主成分1	主成分2	主成分3	象限
y1/所得	0.93	-0.18	-0.24	4
y1/客室	0.86	0.03	-0.38	1
y1/温泉	0.93	-0.31	-0.13	4
y1/施設	0.89	-0.41	-0.18	4
y2/所得	0.84	0.49	0.03	1
y2/客室	0.71	0.67	-0.08	1
y2/温泉	0.73	0.25	0.00	1
y2/施設	0.95	-0.21	-0.12	4
y3/所得	0.81	-0.13	0.54	4
y3/客室	0.80	0.01	0.54	1
y3/温泉	0.91	-0.29	0.24	4
y3/施設	0.91	-0.39	0.03	4
SCORE	0.77	0.60	0.11	1
SCORE2	0.91	0.19	-0.26	4

### 6.3 重回帰分析

図4は、元データで効率値を変数選択したものである。AICが極小値になる3変数の県民所得、客室、県外観光客数が選ばれた。決定係数は0.95で悪くない。

あてはめの要約	
R2乗	0.95092
自由度調整R2乗	0.947495
誤差の標準偏差(RMSE)	0.060208
Yの平均	0.330385
オブザベーション(または重みの合計)	47

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	3	3.0200894	1.00670	277.7049
誤差	43	0.1558775	0.00363	p値(Prob>F)
全体(修正済み)	46	3.1759669		<.0001*

パラメータ推定値						
項	推定値	標準誤差	t値	p値(Prob> t )	下側95%	上側95%
切片	-0.138814	0.054981	-2.52	0.0153*	-0.249694	-0.027934
県民所得	0.0894206	0.022404	3.99	0.0003*	0.0442382	0.1346029
客室	-0.002073	0.001168	-1.78	0.0829	-0.004429	0.0002821
県外	0.1043606	0.003712	28.11	<.0001*	0.0968738	0.1118474

図4 元データで効率値を変数選択

図5は、元データで非効率値を変数選択したものである。AICが極小値になる4変数の客室、温泉地数、県内と県外観光客数が選ばれた。決定係数は0.95で悪くない。ただし、[13]の分析結果に比べ有用な考察は得にくいようだ。

あてはめの要約	
R2乗	0.952785
自由度調整R2乗	0.948289
誤差の標準偏差(RMSE)	1.028548
Yの平均	4.742671
オブザベーション(または重みの合計)	47

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	4	896.63663	224.159	211.8885
誤差	42	44.43227	1.058	p値(Prob>F)
全体(修正済み)	46	941.06890		<.0001*

パラメータ推定値						
項	推定値	標準誤差	t値	p値(Prob> t )	下側95%	上側95%
切片	2.966485	0.465849	6.37	<.0001*	2.0263634	3.9066065
客室	-0.11512	0.017744	-6.49	<.0001*	-0.15093	-0.07931
温泉地	-3.697336	1.155557	-3.20	0.0026*	-6.029344	-1.365328
県内	3.1338155	0.219483	14.28	<.0001*	2.6908809	3.57675
県外	0.7751412	0.085327	9.08	<.0001*	0.6029453	0.9473371

図5 元データで非効率値を変数選択

図6は、比データで効率値を変数選択したものである。AICが極小値になる6変数のy1/温泉地, y1/観光施設, y2/所得, y2/客室, y3/客室y3/観光施設が選ばれた。決定係数は0.99で悪くない。

あてはめの要約						
R2乗					0.992545	
自由度調整R2乗					0.991207	
誤差の標準偏差(RMSE)					0.024639	
Yの平均					0.330385	
オブザベーション(または重みの合計)					47	
分散分析						
要因	自由度	平方和	平均平方	F値		
モデル	7	3.1522899	0.450327	741.7669		
誤差	39	0.0236769	0.000607	p値(Prob>F)		
全体(修正済み)	46	3.1759669		<.0001*		
パラメータ推定値						
項	推定値	標準誤差	t値	p値(Prob> t )	下側95%	上側95%
切片	0.0061344	0.006414	0.96	0.3447	-0.006838	0.0191074
y1/温泉	0.0003972	8.328e-5	4.77	<.0001*	0.0002287	0.0005656
y1/施設	0.0016293	0.000769	2.12	0.0404*	0.0000746	0.0031839
y2/所得	-0.055156	0.016821	-3.28	0.0022*	-0.089179	-0.021132
y2/客室	2.6498138	0.139869	18.94	<.0001*	2.3669025	2.9327252
y3/所得	1.6591799	0.144699	11.47	<.0001*	1.3664982	1.9518617
y3/客室	-6.006947	1.743492	-3.45	0.0014*	-9.533492	-2.480402
y3/施設	-0.022963	0.002657	-8.64	<.0001*	-0.028337	-0.017589

図6 比データで効率値を変数選択

図7は、比データで非効率値を変数選択したものである。AICが極小値になる4変数のy1/所得, y1/客室, y2/所得, y2/客室が選ばれた。決定係数は0.95で悪くない。

あてはめの要約						
R2乗					0.957232	
自由度調整R2乗					0.953159	
誤差の標準偏差(RMSE)					0.978916	
Yの平均					4.742671	
オブザベーション(または重みの合計)					47	
分散分析						
要因	自由度	平方和	平均平方	F値		
モデル	4	900.82128	225.205	235.0108		
誤差	42	40.24762	0.958	p値(Prob>F)		
全体(修正済み)	46	941.06890		<.0001*		
パラメータ推定値						
項	推定値	標準誤差	t値	p値(Prob> t )	下側95%	上側95%
切片	-0.132264	0.22996	-0.58	0.5683	-0.596341	0.3318139
y1/所得	7.7353679	2.22184	3.48	0.0012*	3.2515133	12.219223
y1/客室	9.8532483	24.64545	0.40	0.6913	-39.88328	59.589775
y2/所得	-1.221007	0.686511	-1.78	0.0825	-2.606441	0.1644276
y2/客室	29.593068	6.947271	4.26	0.0001*	15.572907	43.613229

図7 比データで非効率値を変数選択



本節は比データが元データより SCORE と SCORE2 をよく予測できるかを多くの事例で調べる目的で行っている。今後別途報告したい。

## 7. まとめ

本研究では平井 [15] のデータを用いて、47都道府県の宿泊型の観光分析を行った。分析結果から滞在型の観光では圧倒的に北海道が他府県と異質であり、北海道は $\text{km}^2$ あたりの温泉地や観光地が少ないのに DEA 効率的であることが分かった。これは県内観光客といっても面積の大きな北海道は4-5県分の面積がありかさ上げされる。同じ比較をするのは道州制ぐらいにまとめて分析しないとこの問題点は解消できない。しかし九州を比較に取り上げるとしても温泉地や観光施設は九州のほうが多いであろう。結局のところ、北のイメージがうまく働いているのかもしれない。

今後の研究としては、日帰り観光あるいは博物館や美術館あるいはテーマパークという個別テーマで分析することが考えられる。この場合は、すでに良い成果が得られている東京都23区の公立図書館の分析 [11] が役に立つと考える。本年度はローマでの DATA Analytics2014 から8月30日に帰国し、31日から奈良教育大学で行われた「基盤研究 (A) : 非対称・非線形統計理論と経済・生体科学への応用」で本研究の一部を発表した。筆者に先立って、松江工業高等専門学校の村上先生が、観光客に対してアンケート調査を行った結果を報告された [17]。このようなお金も時間もかかる分析も重要であるが、それらを下支えする、山陰の温泉地と観光施設の DEA 分析の共同研究を提案した。一応前向きで検討する旨の回答をいただいた。

また本研究で分かった問題点は、本来平井先生がデータを集めなおして再分析し、論文にまとめられるのが筋であろう。連絡先を調べたが分からないので、本論文を読まれた方で平井先生の連絡先をご存知の方は連絡いただきたい。

(成蹊大学経済学部特任教授)

## 参考文献

- [1] J. P. Sall, L. Creighton & A. Lehman (2004). JMP を用いた統計およびデータ分析入門 (第3版). SAS Institute Japan (株). [新村秀一監修].
- [2] L. Schrage (2003). *Optimizer Modeling with LINGO*. LINDO Systems Inc.
- [3] S. Shinmura (2014). End of Discriminant Functions Based on Variance Covariance Matrices. *2014 ICORE*, 5-16.
- [4] \_\_\_\_\_ (2014). Comparison of Linear Discriminant Functions by K-fold Cross Validation.

*Data Analytics 2014*, 1-6.

- [5] \_\_\_\_\_ (2014). Three Serious Problems and New Facts of the Discriminant Analysis. 1-16. Springのレクチャーシリーズの叢書に収録予定.
- [6] 新村秀一 (2004). JMP活用 統計学とおき勉強法. 講談社.
- [7] \_\_\_\_\_ (2007). ExcelとLINGOで学ぶ数理計画法. 丸善.
- [8] \_\_\_\_\_ (2011). 数理計画法による問題解決法. 日科技連出版社.
- [9] \_\_\_\_\_ (2011). DEAによる回帰型データのクラスター分析, 成蹊大学一般研究報告, 45/3, 1-37.
- [10] \_\_\_\_\_ (2012). 「SAS/JMPとの歩み」. STN (SAS Technical News) 春号~冬号.
- [11] \_\_\_\_\_ (2013). DEA利用のための実践的な解説書. 成蹊大学経済学部論集, 44/1, 15-41.
- [12] \_\_\_\_\_ (2013). 製品の価格性能比の考察—日本車44車種のDEA法と統計手法による分析—. 成蹊大学経済学部論集, 44/2, 1-30.
- [13] \_\_\_\_\_ (2014). 日本の金融機関20行の分析—DEAと統計分析の標準化—. 成蹊大学経済学部論集, 45/1, 33-69.
- [14] 杉山学 (2010). 経営効率分析のためのDEAとinverted DEA. 静岡学術出版.
- [15] 刀根薫 (1993). 経営効率性の測定と改善—抱絡分析法DEAによる—. 日科技連出版社.
- [16] 平井貴幸 (2010). 観光客を「効率的」に誘致している都道府県を探る—DEAによる効率性分析—. 地域と経済, 7. 111-116.
- [17] 村上享 (2014). 中海・宍道湖・大仙圏域における観光動向調査に関する分析—観光統計を活用した地方の観光構造に関する空間分析を中心として—. 「基盤研究(A):非対称・非線形統計理論と経済・生体科学への応用」. 1-8.

DEAに関する論文は成蹊大学レポジトリ (<http://repository.seikei.ac.jp/>) に論文が掲載されています。