

日本の金融機関20行の分析

— DEAと統計分析の標準化 —

新 村 秀 一

1. はじめに

DEAを重回帰分析と比較し、1) 多入力多出力が扱え、2) 求める重みが個々の評価対象ごとに最適な重みを求める点が共通の重みを回帰係数などで求める統計手法に比べて優れている、というような単純な比較が行われている。しかし、重回帰分析に比較し、扱う入力（説明）変数と出力（目的）変数が原則正の値であることが望ましいという扱いにくさや、用いる変数の選択法がないなどの問題もある。これに加えて、数理計画法の歴史的伝統で定式化したモデルを双対モデル¹に置き換えて議論するため、数理計画法の授業が諸外国に比べ劣っている日本において、企業における経営効率性をうたっているDEAの普及を難しいものになっている。筆者はDEAを企業で普及するためには、理系の優秀な大学院生や教員レベルでなく本学部の学生が理解し自由に使えるこなせなければ企業で普及することはおぼつかないと考え、次のような方策を実行し提言してきた。

- 1) DEAの入門手法であるCCRモデルと逆CCR (Inverted CCR) モデルを扱う汎用モデルをLINGO [2] [8] で定式化した。企業で利用し普及するのは、これらの入門モデルで十分と考える。学生が企業などの経営資源と経営の成果を表す入出力データをExcel上にセル範囲名Fと定義するだけで、個々の評価対象 (DMU) に最適な重みWと、その重みを用いて計算した全評価対象のクロス効率値Sと、その対角要素で、各評価対象に最適な効率値SCOREをExcel上に出力する。学生は、各評価対象のDEA効率値が1になるもの (DEA効率率的) とならないものに分けて考えることで問題点が明らかになる。
- 2) DEAを総合化された入力と出力の比を最大化する重みを求め、その重みを全評価対象 (Decision Making Unit, DMU) に適用してクロス効率値を求めて1以下にするという制約を課している。これによって、自分に最適な重みを求めたにもかかわらず、他の対象が1になるためにDEA効率値が1未満になるDMUが現れる。この場合は、与えたデータから定量的にどこを改善すればよいかは筆者の開発した「1入力固定改善法」で簡単に分かる。

¹ 数理計画法 (Mathematical Programming, MP) ソルバーの計算能力が劣っていた時代、双対問題を考えて制約式数と変数の数を逆転することで計算時間が早くなった。また非線形計画法 (Non Linear Programming, NLP) よりLPの方が早いで、DEA法は分数計画法で定式化されるがLPに変形したモデルを扱ってきた。しかし、MPソルバーが早くなったので、双対問題に置き換える事やLPに変換しないで直接NLPで扱うこともできる。本研究では双対問題変換しないLINGOモデルで計算している。

これによって評価が可視化と公平性が達成される。

- 3) DEA 効率的なDMUの作るDEA フロントアで全てのDMUが包括される。そして原点から各DMUを通る直線とDEA フロントアの交点が理想的な改善目標であることが1入力2出力あるいは2入力1出力モデルの比の散布図で説明される。しかしこの説明は2入力2出力以上で説明ができない。これを解決するための拡張解釈として、クロス効率値からDEA クラスタ[3]という概念を導入した。しかしDEA クラスタ毎に改善目標が異なると、改善活動にならないのではないかと考えられる。また、CCRモデルでDEA 効率的であっても、それらの間に優劣があるはずで、この点が従来のDEA分析では考慮されていない。
- 4) p 入力 q 出力モデルでは、 $p \times q$ 個の1入力1出力の比ができて、個々の比の解釈は容易である。しかし、比が最大になるDMUが最大 $p \times q$ 個出てくる可能性がある。この点は企業などで経営改善を行う場合に問題になる。そこでCCRモデルでDEA 効率的になるDMUの中から、逆CCRモデルでDEA 非効率値が最大になるDMUを取りあえず改善活動の最初の目標とし、「1入力固定改善法」で改善案を考えることを指導している。

以上に加えて、本年度からはDEAと統計手法のコラボレーションを実施している。DEAで得られた結果を仮説と考えて、統計手法で仮説検証することである。すなわちCCRモデルで効率的になったDMUと逆CCRモデルで非効率になったDMUを統計手法でベンチマークにして検証できる。またDEA分析は、データ F から、最適な重み W とクロス効率値 S とDEA効率値SCOREが求まる。これらには $S = F * W^T$ と $SCORE = \text{Trace}(S)$ という関係がある。統計分析の狭い世界で考えれば、 F だけが分析対象である。しかし、因果関係のあるデータをDEAで分析すれば、 F に加えて、 $W, S, SCORE$ という新しいデータを分析することで多くの結果が得られる。さらにDEAは比データを扱うので、 $p \times q$ 個の比データも F に加えて分析の対象にできる他、逆CCRモデルで同量のデータが付加される。本研究では、DEAの分析結果を、統計手法で仮説検証する方法を提案する。

2. 分析に用いるデータ (20銀行の6入力2出力モデル)

2013年度の卒業ゼミの江頭さんは、金融業の効率性分析をテーマに選んだ。当初、就活に関係した16行と、表1のデータの従業員数(万人)、本支店数(百ヵ所)、経常費用(兆円)、資本金(兆円)を4入力とし、預金残高(百兆円)、貸出金(十兆)、経常利益(千億円)、純利益(千億円)を4出力とする、4入力4出力で分析した結果を報告した。そこで、メガバンク3行、信託銀行3行、ネット銀行3行、地方銀行3行、信金4行は良い選択であるが、横浜銀行などの地銀上位銀行を加えて20行にすることをアドバイスした。20の制限は、JMPの評価版[1]で登録済みのデータ以外の外部データは、読み込んだ最初の20件しか分析できないためである。学生には、DEAによる分析は、必ず統計分析の結果とあわせて分析するよう

に2013年から次のように指導方針を変えた。DEA分析の標準化は頁数の関係で省略する。統計分析の標準化は4章以降の統計分析を実施すればよい。

- 1) 従来の統計分析では比率情報の分析が軽視されてきた。因果関係を分析する重回帰分析や判別分析では、原因と考える説明変数を入力とし、結果の目的変数を出力と考える。因果関係を表す1入力1出力の比を作成し、元データに加えて比データの統計分析を検討することを課した。4入力4出力の分析では、16個の比率尺度が得られる。比率は単純で分かり易いが、各DMUそれぞれが16個全ての比率で一番良くなる可能性がある。これは、入出力の変数の個数の p と q が増えると、DEA効率的なDMUが増える原因になる。一般的にDMUが n 個とすれば、 $n \geq \text{MAX}(p \cdot q, 3 \cdot (p + q))$ であることが良いとされている。“ $p \cdot q$ ”は上で述べたことに対応しているが、 $3 \cdot (p + q)$ の3は多分経験的なものであろう。この式から $\text{MAX}(p \cdot q, 3 \cdot (p + q)) = \text{MAX}(16, 24) = 24$ 行以上の銀行が望ましいことになる。しかし、CCRモデルでDEA効率値が1になるDMUが多いことがDEAにとって問題であるか疑問である。電力業界や金融業界や病院のように規制がある場合、効率的なDMUが他の業種より多いことが考えられ、広く比較することで業種比較が行えると考えられる。
- 2) 最初に学生が行った4入力4出力モデルの分析結果は、16行の金融機関のうち10行が効率的になった。これは「DMU数が少ないためと考えられるが、効率値の最小値は0.72と上級演習で行った東京都の公立図書館と異なり、規制がある業界なので効率性に大きな影響のないためと考えられる」とコメントした。このような裏づけのないコメントは間違っている可能性がある。幸い、2010年に最適線形判別関数の基礎研究が終了し、2012年には応用研究が終了したので、学生の分析を間違いなく指導するため自分で学生とは独立して分析を先回りして行い指導することにした。また入出力項目に漏れがないかを幸村先生に相談し、国債と海外資産を入力に加えることのアドバイスを受けた。しかし「今回は国内の本支店と従業員に限定しているので、海外資産を入れることはおかしい」というので国債残高²と預金残高を入力として扱うことをアドバイスした。本人も納得してまとめたのが表1の6入力3出力モデルである。DMUは地銀4行を加えた20行で、上の式の27行以上が望ましいことに反するが、これで行うことを基本方針にした。
- 3) 井上智雄先生から金融分野で行われているDEA分析の代表的な論文を紹介してもらったが、研究論文であり、学生には荷が重すぎた。しかし、最後のまとめに「本来であれば、1金融機関の本支店の分析のほうが良い分析になるだろう」というコメントは、私が指導学生にも社会人になって、実際にDEAを利用する際に薦めていることである。さらに付け加えて必ず事前に次の点を根回しておく必要を徹底している。すなわち、「DEAで効率

² 江頭の6入力3出力の分析では、国債と地方債を合算したものを扱うべきであった事が分かった。

的になるのは、経営資源の使用の割りに出力で頑張っているDMUを探し、新しい問題点を検討し関係者全員で改良することが目的である。企業の事業部評価を行う場合、No.1の事業部が効率的になるとは限らないことを事前に言うておかないと、将来トップになるであろう事業本部長の理解が得られない。それよりも、各事業部に属する部の評価を行えば、各事業部長の協力を得られやすいからである。すなわち、品質管理が製造現場の改善活動であれば、DEAは「評価の可視化と公平性」から、ホワイトカラーが直面する問題点の発見が容易であり、改善点を話し合う「評価と改善のための手法」として用いるべきだと考えている。自分に最適な重みで評価したが、1以下の場合には自分の重みで他のDMUが1になるので、数量的にどこが悪いかわかる。ただし、入出力の項目の違いによって、著しく結果が変わる場合もあるので、どこが問題点かを時間をかけて関係者全員で検討し、より良いものに成長させていく必要がある。

- 4) さらに分析上の注意点であるが、単位をうまく選んで最大値が1以上10未満になるように指導している。このため表1の金額の単位が異なっている。このようにしないと、「最適化計算の途中で大きな値で小さな値を割って0判定されれば、それ以降の計算に0の値が用いられる数値計算上のトラブル」を避けるためである。また、計算された重みの解釈はスケールが揃っていると容易になる。

表1 分析に用いる6入力3出力モデルのデータ

SN	銀行	種別	従業員数 (万人)	本支店数 (100)	経常費用 (兆円)	資本金 (兆円)	国債保有額 (十兆円)	預金残高 (百兆円)	貸出金 (十兆)	経常利益 (千億円)	純利益 (同左)
1	三菱東京	1	3.65	6.72	1.94	1.71	4.18	1.18	7.41	8.61	5.85
2	三井住友	1	2.26	5.05	1.45	1.77	2.62	0.80	5.98	6.71	6.18
3	みずほ銀行	1	2.66	4.20	1.73	1.40	4.58	0.95	6.35	5.35	4.85
4	みずほ信託	2	0.31	0.36	0.14	0.02	0.09	0.03	0.37	0.35	0.26
5	三菱UFJ信託銀行	2	0.69	0.59	0.39	0.32	0.68	0.12	1.13	1.36	1.25
6	三井住友信託	2	1.28	1.20	0.53	0.34	0.24	0.23	2.23	7.25	1.05
7	オリックス	3	0.05	0.01	0.02	0.05	0.01	0.01	0.92	0.10	0.05
8	セブン銀行	3	0.05	0.01	0.06	0.03	0.01	0.00	0.00	0.32	0.20
9	ソニー銀行	3	0.04	0.01	0.03	0.03	0.02	0.02	0.10	0.04	0.01
10	東京都民銀行	4	0.16	0.71	0.04	0.05	0.04	0.02	0.18	0.02	0.02
11	八千代銀行	4	0.16	0.79	0.03	0.04	0.03	0.02	0.14	0.07	0.05
12	東京スター銀行	4	0.12	0.31	0.06	0.03	0.03	0.02	0.15	0.07	0.02
13	横浜銀行	4	0.39	1.96	0.16	0.22	0.08	0.11	1.15	0.86	0.53
14	千葉銀行	4	0.40	1.73	0.13	0.15	0.11	0.10	0.79	0.66	0.44
15	福岡銀行	4	0.35	1.62	0.12	0.08	0.10	0.08	0.69	0.66	0.33
16	静岡銀行	4	0.29	1.71	0.11	0.09	0.13	0.08	0.70	0.65	0.41
17	多摩信金	5	0.20	0.77	0.04	0.02	0.01	0.02	0.10	0.05	0.04
18	城北信金	5	0.21	0.81	0.04	0.03	0.02	0.02	0.11	0.07	0.06
19	巣鴨信金	5	0.12	0.43	0.02	0.00	0.01	0.02	0.08	0.04	0.03
20	城南信金	5	0.21	0.85	0.04	0.05	0.03	0.03	0.19	0.50	0.05

以上をまとめれば、DEA分析と統計分析のコラボレーションとして次の事を提案する。

- 1) 入力データFだけの統計分析では効果が乏しい。
- 2) しかし、DEAで事前に分析すれば、DEAの分析結果を統計で検証することになり、分析がしやすくなる。また、Fから重みWとクロス効率値Sの関係が $S=F*W^T$ で得られ、効率値SCOREが $SCORE=Trace(S)$ になる。さらに逆CCRモデルで同量の情報が得られ、分析すべきデータが格段に増える。
- 3) DEA分析を入力と出力の比の情報を分析する手法と考えれば、回帰分析や判別分析のような因果関係にあるデータは、入力と出力の比データを分析に加えることが考えられる。

3. 6入力2モデルの分析

3.1 CCRモデルによる分析

表2はCCRモデルによるDEA効率値(SCORE)と重み(入力:W1-W6,出力:W7,W8)である。江頭さんには6入力3出力モデルの分析を指導したが、筆者は経常費用があるので純利益を出力から省いた6入力2出力モデルで分析し、後で違いを比較評価することにした。また、学生を指導する場合、学生の分析結果と同じであれば教育効果がないと考えた。さらにDEAには変数選択法(今回は2出力なので、統計的にも変数選択は難しい)がなく、似た項目の影響が分からないことである。DEAは統計より優れた点として、多くのDEA研究で多入力多出力を扱える点を上げている。しかし日本車40車種のDEA分析に見るとおり[5]、4性能を入力として販売台数と販売金額を2出力とするよりも、販売台数と販売金額を個別に分析した方がDEA効率的なDMUも少なくなり、解釈が分かりやすくなった。

6入力2出力で分析すると、表2の5行が効率的になった³。みずほ信託と三井住友信託の信託2行、オリックスとセブン銀行のネット銀行2行、巢鴨信金1行が選ばれた。預金残高を出力にした結果と異なり、いくら規制されていても競争の激しい業界であり東京都民銀行のように効率値が0.18のものが現れるのは常識的に妥当である。また金融業としての業務形態(種別)や規模が異なり各種別から信託銀行、ネット銀行と信金の代表行が選ばれた。ゼミでは、取りあえず効率的な5行が選ばれた理由をこれまでの金融に関する授業で習った知識を踏まえて議論させ、それが正しいかを調べる次のような課題を与えた。

- 1) メガバンクはなぜ選ばれず、信託銀行はなぜ2行選ばれたか?
- 2) ネット銀行でオリックスが選ばれたのは不動産などの高額融資とATM等の設備投資が少ないこと、セブン銀行はATMでリテールを行っているとの発表があった。他の銀行で

³ 江頭の6入力3出力ではメガバンクから三井住友、地方銀行から横浜と静岡の3行が加わり8行が選ばれた。銀行は規模や形態が種々あるが各種別から手本が選ばれた。

ATMを担当する部署があれば、将来的には他の銀行はATM部門を独立させて従来部門やインターネット銀行と競わせることも考えられる。

3) 地方銀行がメガバンク同様選ばれなかった理由。

4) 巢鴨信金はマスコミでもよく取り上げられるが、具体的にどの経営指標が良いのかを「一入力固定改善法」で明らかにする。

重みの分析までは、これまで学生には求めていなかった。DEA研究では、重みが0になるのは良くないと指摘するものもある。確かに都合の悪い入出力項目を0の重みで無視することで効率値の最大化を図っている。しかしメガバンク3行で比較すれば、三菱東京と三井住友銀行は従業員数と経常費用が正の重みで、本支店数、資本金、国債保有額と預金残高の4入力に0である。これらの入力の重みが正であれば効率値を下げる効果があると考えられるので、この2行では他行より大きいこれらの変数の重みを0にしていると考えられる。これに対してみずほ銀行は従業員数と資本金が正の重みであり、他の2行より経常費用が多く資本金が少ないことを表していると考えられる。三井住友信託は経常費用と預金残高が少ない割に出力が多いと考えられる。オリックスは本支店数と経常費用が少なく、セブン銀行は本支店数が少ないことが分かる。セブン銀行はATMの設備投資や運用に費用がかかる点が、オリックスと異なっているようだ。オリックス銀行は、インターネットで不動産融資など高額商品を扱っているとの報告があったが、インターネット銀行3行をより詳しく調査してもらうことにした。八千代銀行と横浜銀行と千葉銀行と城南信金は経常費用、巢鴨信金は資本金が少ない割に出力が良いことが読み取れる。すなわちDEAは、悪い点に目をつぶり良い点を評価する手法であり、0を忌避する理由が分からない。

表2 CCRモデルによるDEA効率値と重み

SN	銀行	種別	SCORE	W1	W2	W3	W4	W5	W6	W7	W8
1	三菱東京UFJ銀行	1	0.456	0.249	0	0.047	0	0	0	0.01	0.044
2	三井住友銀行	1	0.567	0.396	0	0.074	0	0	0	0.016	0.07
3	みずほ銀行	1	0.405	0.314	0	0	0.119	0	0	0.017	0.056
4	みずほ信託銀行	2	1	0	0.126	0	38.61	0	0	1.354	1.424
5	三菱UFJ信託銀行	2	0.377	1.312	0	0.245	0	0	0	0.053	0.233
6	三井住友信託銀行	2	1	0	0	0.509	0	0	3.181	0	0.138
7	オリックス銀行	3	1	0	29.36	31.88	0	0	0	0.345	7.099
8	セブン銀行	3	1	0	100	0	0	0	0	0.758	3.123
9	ソニー銀行	3	0.292	0	25.39	27.57	0	0	0	0.299	6.139
10	東京都民銀行	4	0.179	0	0	0	19.52	0	2.585	0.915	0.634
11	八千代銀行	4	0.207	0	0	28.75	0	0	0	0.485	1.97
12	東京スター銀行	4	0.334	0.657	0	0	35.43	0	0	1.628	1.285
13	横浜銀行	4	0.486	0	0	6.18	0	0	0	0.104	0.424
14	千葉銀行	4	0.451	0	0	7.668	0	0	0	0.129	0.525
15	福岡銀行	4	0.625	0.209	0	0	11.25	0	0	0.517	0.408

16 静岡銀行	4	0.58	0	0	1.117	9.707	0	0	0.458	0.399
17 多摩信用金庫	5	0.264	0	0	0	38.1	7.647	0	1.794	1.496
18 城北信用金庫	5	0.232	0	0	3.26	28.33	0	0	1.337	1.164
19 巣鴨信用金庫	5	1	0	0	0	347.9	0	0	5.073	14.84
20 城南信用金庫	5	0.9	0	0	24.06	0	0	0	0.406	1.649

表3はクロス効率値である。S1列からS20列は20銀行の重みで、20銀行の効率値をSN = 1から20まで効率値を計算し示す。SN=4のS1列からS20列は20行の重みによるみずほ信託の効率値である。自行と東京都民銀行の重みで1になる。すなわち、CCRモデルで効率的であるがあまり影響力がないことを示す。SN=6は三井住友信託の20行の重みによる効率値で、セブン銀行と東京都民銀行以外の18行の重みで効率的である。SN=7（オリックス銀行の効率値）は、S4（みずほ信託）とS6（三井住友信託）とS19（巣鴨信託）だけで1にならず、残り17行では効率的になっている。SN=19の巣鴨信託は、S4（みずほ信託）、S10（東京都民銀行）、S12（八千代銀行）、S15からS19（千葉銀行から巣鴨信託）の重みで効率値が1になっている。以上から、逆CCRモデルで逆SCOREを見なくても三井住友信託かオリックス銀行を手本に「1入力固定改善法」で改善案を検討することが考えられる。

表3 クロス効率値

SN	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
1	0.456	0.456	0.449	0.333	0.456	0.25	0.246	0.048	0.246	0.335	0.369	0.367	0.369	0.369	0.367	0.364	0.269	0.364	0.278	0.369
2	0.567	0.567	0.517	0.256	0.567	0.282	0.255	0.05	0.255	0.265	0.386	0.286	0.386	0.386	0.286	0.288	0.237	0.288	0.211	0.386
3	0.404	0.404	0.405	0.296	0.404	0.189	0.225	0.051	0.225	0.308	0.273	0.334	0.273	0.273	0.334	0.324	0.219	0.324	0.228	0.273
4	0.228	0.228	0.255	1	0.228	0.29	0.175	0.038	0.175	1	0.223	0.975	0.223	0.223	0.975	0.791	0.72	0.791	0.82	0.223
5	0.377	0.377	0.373	0.275	0.377	0.327	0.337	0.087	0.337	0.286	0.286	0.3	0.286	0.286	0.3	0.296	0.232	0.296	0.23	0.286
6	1	1	1	1	1	1	1	0.203	1	0.914	1	1	1	1	1	1	1	1	1	1
7	1	1	1	0.798	1	0.293	1	1	1	1	1	1	1	1	1	1	1	1	0.39	1
8	1	1	1	0.387	1	1	1	1	1	0.335	0.353	0.371	0.353	0.353	0.371	0.35	0.394	0.35	0.448	0.353
9	0.273	0.273	0.263	0.161	0.273	0.081	0.292	0.207	0.292	0.177	0.169	0.19	0.169	0.169	0.19	0.186	0.179	0.186	0.105	0.169
10	0.07	0.07	0.081	0.142	0.07	0.035	0.011	0.003	0.011	0.179	0.108	0.178	0.108	0.108	0.178	0.178	0.168	0.178	0.076	0.108
11	0.107	0.107	0.111	0.161	0.107	0.118	0.023	0.004	0.023	0.189	0.207	0.19	0.207	0.207	0.19	0.197	0.184	0.197	0.115	0.207
12	0.139	0.139	0.156	0.289	0.139	0.096	0.049	0.011	0.049	0.323	0.125	0.334	0.125	0.125	0.334	0.305	0.301	0.305	0.196	0.125
13	0.477	0.477	0.459	0.325	0.477	0.266	0.104	0.018	0.104	0.355	0.486	0.377	0.486	0.486	0.377	0.383	0.381	0.383	0.249	0.486
14	0.351	0.351	0.351	0.347	0.351	0.239	0.091	0.015	0.091	0.37	0.451	0.396	0.451	0.451	0.396	0.404	0.379	0.404	0.275	0.451
15	0.384	0.384	0.399	0.553	0.384	0.286	0.096	0.016	0.096	0.577	0.476	0.625	0.476	0.476	0.625	0.62	0.573	0.62	0.463	0.476
16	0.468	0.468	0.477	0.504	0.468	0.294	0.091	0.015	0.091	0.532	0.534	0.58	0.534	0.534	0.58	0.58	0.499	0.58	0.418	0.534
17	0.068	0.068	0.074	0.214	0.068	0.081	0.018	0.003	0.018	0.245	0.152	0.245	0.152	0.152	0.245	0.254	0.264	0.254	0.162	0.152
18	0.078	0.078	0.083	0.192	0.078	0.1	0.021	0.004	0.021	0.221	0.182	0.22	0.182	0.182	0.22	0.232	0.23	0.232	0.146	0.182
19	0.083	0.083	0.094	1	0.083	0.088	0.023	0.004	0.023	1	0.186	1	0.186	0.186	1	1	1	1	1	0.186
20	0.434	0.434	0.421	0.437	0.434	0.526	0.137	0.02	0.137	0.425	0.9	0.458	0.9	0.9	0.458	0.496	0.476	0.496	0.439	0.9

3.2 逆CCRモデル (Inverted CCRモデル [6] [7])⁴

表4は、逆CCRモデルの結果である。オリックス銀行の逆効率値が最大の4.19で、次の三井住友信託銀行は2.49である。筆者は、取りあえずCCRモデルで効率的で、Inverted CCRモデルで逆効率値が最大なDMUの構成比を参考に「1入力固定改善法」で改善策を考えるように提案した [4]。

出力の重みは、オリックス銀行は入力が資本金、出力は経常利益だけが正である。すなわち資本金の割に経常利益が少ないことを示す。

表4 逆CCRモデルの逆効率値と重み

SN	銀行	種別	SCORE	W1	W2	W3	W4	W5	W6	W7	W8
1	三菱東京UFJ銀行	1	1.076	0	0	0	0	0.083	0.552	0.135	0.008
2	三井住友銀行	1	1.072	0	0	0	0.403	0.109	0	0.134	0.041
3	みずほ銀行	1	1	0	0	0	0	0.218	0	0.154	0.005
4	みずほ信託銀行	2	1.193	0.354	0	5.093	0	2.168	0	2.191	1.083
5	三菱UFJ信託銀行	2	1.008	0	0	0	1.972	0.534	0	0.654	0.199
6	三井住友信託銀行	2	2.49	0	0	0	0	0.569	3.774	0.926	0.058
7	オリックス銀行	3	4.191	0	0	0	22.22	0	0	0	43.7
8	セブン銀行	3	1	0	0	16.11	0	0	0	4.701	3.119
9	ソニー銀行	3	1	0	0	0	9.606	5.489	32.07	9.68	1.421
10	東京都民銀行	4	1	0	0.024	4.733	12.89	4.172	0	5.29	2.234
11	八千代銀行	4	1	0	0	0	18.92	5.121	0	6.274	1.91
12	東京スター銀行	4	1	0	0	11.63	0	4.185	9.179	5.499	2.456
13	横浜銀行	4	1.802	0	0.024	0	3.133	0	2.422	1.323	0.328
14	千葉銀行	4	1.648	0	0.024	0	2.836	1.117	4.245	1.792	0.346
15	福岡銀行	4	1.882	0	0	0	0	1.584	10.5	2.575	0.162
16	静岡銀行	4	1.831	0	0	0	0	1.521	10.08	2.473	0.155
17	多摩信用金庫	5	1	0	1.299	0	0	0	0	9.829	0.03
18	城北信用金庫	5	1	0	0.884	0	2.633	10.65	0	8.431	0.504
19	巣鴨信用金庫	5	1.066	0	0	0	0	12.82	51.74	8.021	10.64
20	城南信用金庫	5	1.424	0	0	0	0	3.916	25.96	6.367	0.399

杉山氏 [6] に敬意を払い、表2の効率値 (SCORE2の2は出力を2変数にした結果を示す) を横軸に、表4の逆効率値 (逆2) を縦軸にして図1の散布図を描く。杉山氏の提案は、例えばSCORE \geq 0.95で逆SCORE \geq 4のオリックス銀行は非常に効率的であると分類し、筆者の「1入力固定改善法」の改善目標になるDMUと同じになる。しかしオリックス銀行は、インターネット専門銀行に特化しており銀行業全般の目標とするには問題があり、実際には住友信託を目標とすべきであろう。そして逆SCORE \leq 1.25でSCORE \geq 0.95のみずほ信託、セブン銀行、

⁴ 筆者が逆CCRモデルと言ってきたのは、文献 [6] [7] でInverted CCRモデルとして先行研究で提案されていたことが分かった。実はDEAはLINGOのサンプルモデルを実際に使ってみて理解できたので先行研究の調査をきちっと行ってないのだから長く気づかなかった。

巢鴨信金は効率的であるが逆CCRで非効率であり，他行の参考にならない特異的なDEAと考えている。6の住友信託はそれなりに評価できる。SCORE<0.95で1.25 ≤ 逆SCORE ≤ 2.25に布置された横浜銀行から静岡銀行の地銀4行は改善を行い住友信託のレベルに改善することが期待される。SCORE<0.95で逆SCORE ≤ 1.25に布置されたメガバンク3行とその他の7行は経営改善の余地が大きいグループになる。すなわち，図1はDEA効率的なものが全て改善目標になるという従来の立場を，筆者同様に疑問視していると考えられる。

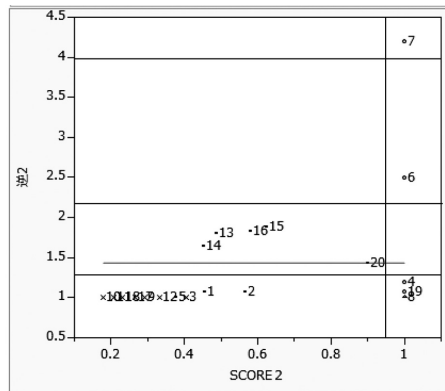


図1 スコアと逆スコアの散布図による分類

表5は，逆CCRモデルのクロス効率値である。SN=9のソニー銀行が非効率であるのは4行に対してであるが，SN8のセブン銀行が18行に対して最も非効率である。これは自己資金でなく他銀行からの借り入れでATMを運用し，リテール業手数料を利用の源泉にしているためと想定される。SN=3のみずほ銀行は16行の重みで非効率でありSN=17の多摩信金は11行に対して非効率である。

表5 逆CCRモデルのクロス効率値

SN	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	
1	1.076	1.17	1.293	1.266	1.17	1.076	9.89	1.979	1.086	1.197	1.17	1.217	1.505	1.106	1.076	1.076	8.375	1.217	1.316	1.076	
2	1.312	1.072	1.659	1.468	1.072	1.312	7.449	2.098	1.181	1.143	1.072	1.402	1.328	1.136	1.312	1.312	8.988	1.451	1.59	1.312	
3	1	1	1	1	1	1	7.489	1.667	1	1	1	1	1.493	1	1	1	11.46	1	1	1	
4	2.181	3.184	2.908	1.193	3.184	2.181	27.67	1.303	2.383	2.029	3.184	1.296	3.804	2.532	2.181	2.181	7.856	2.419	2.423	2.181	
5	1.354	1.008	1.218	1.063	1.008	1.354	8.265	1.508	1.212	1.015	1.008	1.125	1.473	1.139	1.354	1.354	14.51	1.189	1.595	1.354	
6	2.49	3.621	7.219	3.452	3.621	2.49	41.71	3.846	2.676	3.518	3.621	3.229	3.224	2.912	2.49	2.49	14.24	4.989	6.377	2.49	
7	18.49	6.575	57.29	13.71	6.575	18.49	4.191	13.01	10.87	6.964	6.575	13.2	7.509	9.098	18.49	18.49	699.6	31.53	12.12	18.49	
8	1	1	1	1	1	1	20.63	1	1	1	1	1	1	1	1	1	1	1	1	11.57	1
9	1.137	1.007	3.568	1.341	1.007	1.137	2.716	1.353	1	1	1.007	1.127	1	1	1.137	1.137	73.53	2.827	1.02	1.137	
10	1.52	1.059	3.361	1.167	1.059	1.52	1	1.307	1.243	1	1.059	1.187	1.09	1.115	1.52	1.52	1.906	1.315	1	1.52	

11	1.36	1	2.922	1.229	1	1.36	3.205	1.552	1.135	1	1	1.27	1	1.013	1.36	1.36	1.321	1.021	1.246	1.36
12	1.456	1.638	3.304	1	1.638	1.456	5.102	1	1.405	1.27	1.638	1	1.586	1.417	1.456	1.456	3.709	1.899	1.27	1.456
13	2.333	1.98	10.77	3.061	1.98	2.333	7.881	3.106	1.999	2.046	1.98	2.59	1.802	1.913	2.333	2.333	4.445	3.245	2.656	2.333
14	1.755	1.882	5.155	2.343	1.882	1.755	9.007	2.757	1.655	1.896	1.882	2.066	1.716	1.648	1.755	1.755	3.47	2.266	2.039	1.755
15	1.882	2.719	5.161	2.358	2.719	1.882	15.72	2.755	1.947	2.474	2.719	2.135	2.294	2.024	1.882	1.882	3.23	2.293	2.315	1.882
16	1.831	2.354	3.848	2.418	2.354	1.831	14.11	3.123	1.859	2.28	2.354	2.171	2.202	1.89	1.831	1.831	3.104	1.975	2.166	1.831
17	1	1.443	5.48	1	1.443	1	4.582	1.116	1	1.204	1.443	1	1.018	1	1	1	1	1	1	1
18	1.106	1.235	4.331	1.091	1.235	1.106	4.255	1.293	1.042	1.129	1.235	1.116	1	1	1.106	1.106	1.072	1	1.122	1.106
19	1.11	4.927	4.658	1.202	4.927	1.11	27.49	1.409	1.34	2.524	4.927	1.174	2.019	1.545	1.11	1.11	1.403	1.335	1.066	1.11
20	1.424	1.837	5.391	2.779	1.837	1.424	17.9	3.675	1.445	2.057	1.837	2.501	1.519	1.469	1.424	1.424	1.727	1.582	3.217	1.424

3.3 DEAクラスター

表6は表3のクロス効率値から求めたDEAクラスター [3] [9] であり、筆者の提案したものであるがまだ定着していない。1入力2出力モデルあるいは2入力1出力モデルの場合、2個の1入力1出力の比が計算でき、この値で散布図を作れば、効率的なDMUを結んだフロンティアで全てのDMUが内包される[3]。そして原点と隣り合った効率的なDMUとの三角形に含まれるDMUは、原点とDMUを通る直線とその2つのDMUを結ぶ線分上の交点が改善目標になる。このような実態のない理想点を現実の改善活動の目標にすることはできない。また(p+q) が2以上の場合、この散布図で説明できない。それを解消する意味で、DEAクラスターを考えた。D7は6の住友信託、7のオリックス銀行、8のセブン銀行が作る三角形のフロンティアを底辺とし、原点を頂点とする三角錐である。ここに、SN=1,2,3,5,7,9の6行が含まれる。D8は住友信託とオリックス銀行を結ぶ線分がフロンティアになり、原点でできる三角形にSN=8,11,13,14,20の5行がある。D6は住友信託とセブン銀行を結ぶ線分がフロンティアになり、原点でできる三角形にSN=6の1行がある。これらの12行は住友信託、オリックス銀行、セブン銀行のいずれかを手本とすればよいが、ネット銀行は手本になりえないので、住友信託を改善活動の手本とすることが考えられる。D12とD19は19の巣鴨信金、D4とD10は4のみずほ信託と巣鴨信金を手本とすることは適切とは考えられないので、これらを考慮すれば中核となるD7の三角錐を中心にした改善が考えられる。以上の事を行い、必要であれば任意のDMUを手本にして「一入力固定改善法」で改善案の検討を行えばよいことになる。

表6 DEAクラスター

DEAクラスター	手本	数	構成員
D7	6-8	6	1-3,5,7,9
D8	6,7	5	8,11,13,14,20
D6	6,8	1	6
D12	6,7,19	3	12,15-18
D4	4,6,19	1	4
D10	4,7,19	1	10
D19	6,19	1	19

すなわち本データは、D7の3行をフロンティアとする三角錐が中心にあり、みずほ信託と巣鴨信金がそれに隣接する構造と考えられる。

4. 統計分析の標準化

DEA分析で因果関係にあるDMUの種々の情報が得られた。統計分析にこれらの情報を投影して考えれば、これまで以上に因果関係がある分析対象の分かりやすい分析結果が得られると考える。以下の分析は、代表的な統計分析の標準手順である。DMUが20行なので、JMPの評価版が利用できる [1]。

4.1 層別箱髷図

与えられた分析対象が、銀行の種別のように分類されておれば、まず一元配置の分散分析で層別箱髷図を検討することが重要である。分類情報がない場合、重ね合わせプロットや1変数のヒストグラムなどの検討が必要になる。

図2は、SCOREと逆SCOREの出力結果である。みずほ信託と住友信託、オリックス銀行、セブン銀行、巣鴨信金が1である。他の信託銀行とインターネット銀行と城南信金を除く信金は、メガバンクや地方銀行と同じく効率的でないことになる。リテール部門のないオリックス銀行を除くみずほ信託、住友信託、巣鴨信金と城南信金が一般的に考えられる金融機関としての業務を行っていて、効率的と考えられる。逆SCOREではオリックス銀行が、金融業で一番効率的な不動産の融資にインターネットで対応して一番際立っているが、他の銀行と経営効率性を比較するには適していない。信託銀行、地方銀行、信金そしてメガバンクの順に範囲が狭まっている。特に地銀の数行の逆SCOREがメガバンクより大きいことは今後の分析テーマになるかもしれない。

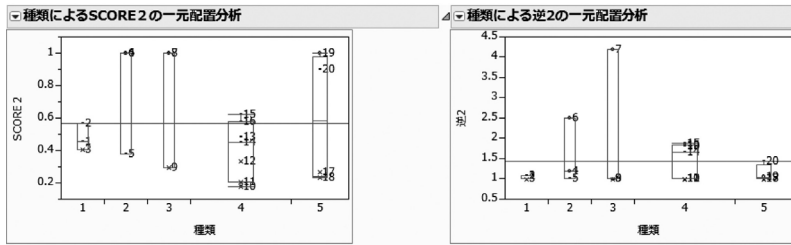


図2 SCOREと逆SCORE

図3は、表1に示す9変数の層別箱髷図である。従業員数は、メガバンク、信託銀行、地方銀行、信金、インターネット銀行の順に少ない。メガバンクの従業員数は1（東京三菱UFJ）、3（三井住友）、2（みずほ）の順であるが、興銀は元々本支店数が少ない事はわかっているが、本支店数は1, 2, 3の順である。従業員数と本支店数が対応していないのは、三井住友銀行とみずほ銀行の合併の前後の従業員数と、本支店数を調べてリストラ効果を検討する必要がある。また地方銀行と信金は、信託銀行に比べ従業員数の割に本支店数が多いのはリテール部門の重要性のためと考えられる。経常費用、資本金、国債保有額、預金残高は、ほぼ従業員数と似た傾向にある。これに対して経常利益は住友信託が際立っているが、これを省けば純利益同様、ほぼ従業員数似た傾向にある。また同社の経常利益と純利益の落差を調べる必要がある。

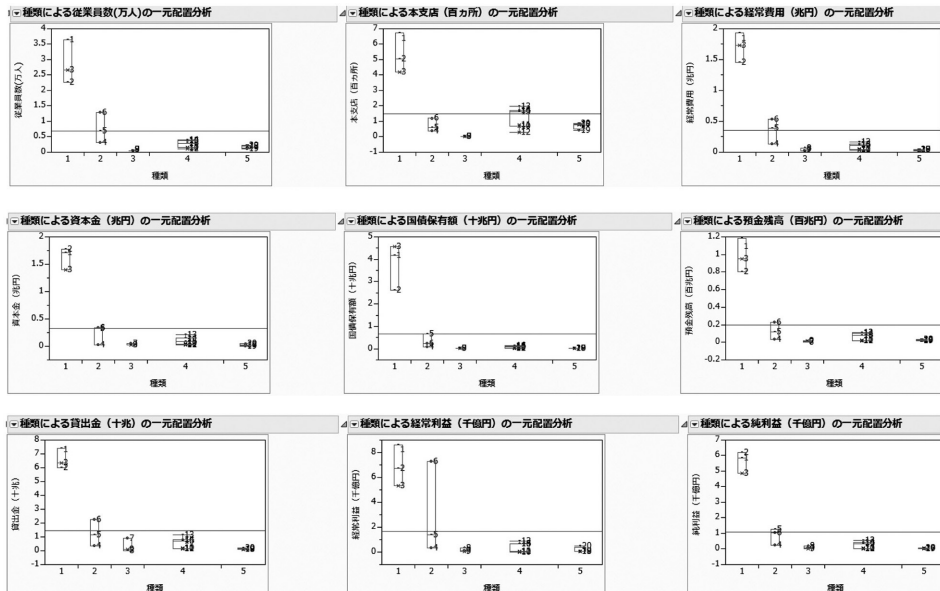


図3 表1に示す9変数の層別箱髷図

図4は、表1に示す9変数から1入力1出力の比を計算し、出力が貸出の6変数の比である。入力が資本金を除く5変数は、オリックスが非常に大きいので、他の銀行の違いが分からない。オリックス銀行を省いた再検討は頁数の関係で省く。資本金に対し、信金、インターネット銀行、信託銀行、地方銀行、メガバンクの順に小さくなるが、みずほ信託とオリックスを省けば、信託銀行とインターネット銀行は地方銀行より少なくなる。

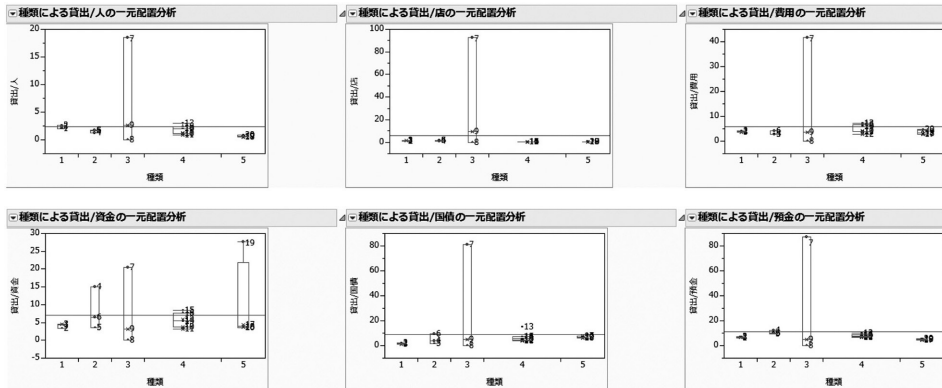


図4 出力が貸出の6変数の比

図5は、出力が経常利益の6変数の比である。住友信託は従業員、経常費用、資本金、国債保有額当たりの経常利益が高いが、本支店数と預金残高当たりの経常費用はそれほど高くない。オリックス銀行は、従業員、本支店数、国債保有、預金残高当たりの経常利益が高いのは業務形態の特異性を表している。これは経常費用、資本金あたりの経常利益が高くないことでも裏付けられる。城南信金は経常費用と国債あたりの経常利益が高く、巣鴨信金は資本金当たりの経常利益が高い。

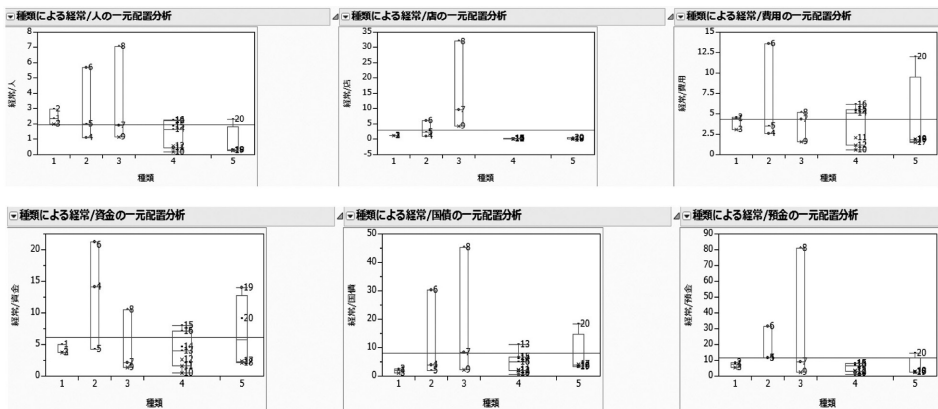


図5 出力が経常の6変数の比

図6は、出力がDEA分析には含めていない純利益の6変数の比である。セブン銀行は、従業員、本支店数、国債、預金残高当たりの純利益が高いが、ATMに稼がせていることから理解できる。このため経常費用、資本金当たりの純益は相対的に低い。メガバンクは、従業員当たりの純益が信託銀行、地方銀行、信金に比べ高いが、これであれば給与水準が高くても吸収できる。経常費用あたりの純益は地方銀行のばらつきが大きい。資本金の純益の比は住友信託、オリックス銀行と巣鴨信金を省くと地方銀行のばらつきが大きい。なぜ住友信託が他の2行と異なるか調べる必要がある。

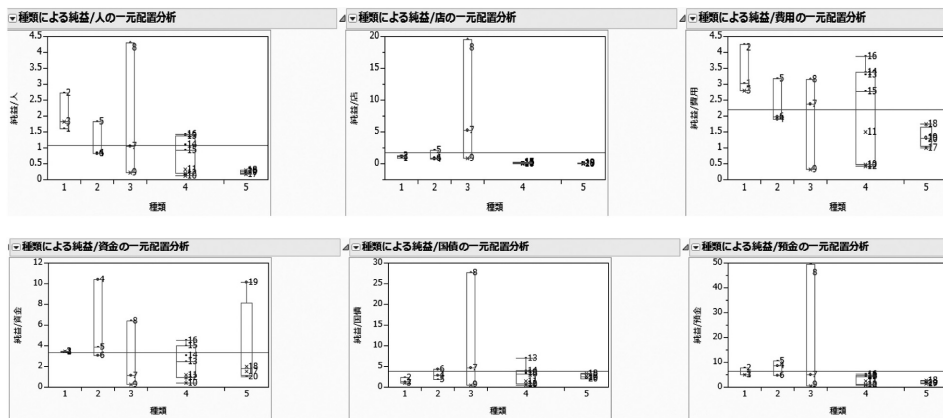


図6 出力が純利益の6変数の比

以上みてきたとおり、DEA分析の結果を箱髭図に投影することで、箱髭図の解釈の視点が明確になる。

4.2 相関行列

表7は、元データの相関行列である。最初の2行はSCOREと逆SCOREとの相関係数で、それ以下は表に示す元データ9変数と貸出金を出力とする6個の比データの相関係数である。SCOREと逆SCOREとは、経常利益と6個の比が正の相関であるが、恐らく逆SCOREと経常利益は無相関であろう。すなわち元データは効率値と負の相関であるが絶対値が0.21より小さい。資本金を除く5個の比と逆SCOREが0.8以上と大きい。従業員数と他の8個の元データの相関は0.927以上と一番相関が高い。これに対して、従業員数、本支店数、国債残高、預金残高を分母とする比は逆SCOREと0.968以上と高いが貸出／資本金と他とは0.444で低い。

表7 元データと貸出との比データの相関行列

従業員 (万人)	本支店 (100)	経常費用 (兆円)	資本金 (兆円)	国債保 (十兆)	預金 残高 (百兆)	貸出金 (十兆)	経常 利益 (千億)	純利益 (千億)	貸出 /人	貸出 /店	貸出/ 費用	貸出/ 資金	貸出/ 国債	貸出/ 預金
-0.08	-0.17	-0.09	-0.11	-0.15	-0.11	-0.06	0.083	-0.1	0.289	0.316	0.305	0.561	0.336	0.334
-0.14	-0.16	-0.17	-0.17	-0.21	-0.17	-0.07	0.003	-0.18	0.862	0.828	0.887	0.438	0.888	0.867
1	0.938	0.989	0.957	0.951	0.988	0.983	0.927	0.954	-0.05	-0.15	-0.14	-0.2	-0.21	-0.13
0.938	1	0.926	0.94	0.89	0.953	0.942	0.83	0.935	-0.08	-0.2	-0.15	-0.23	-0.23	-0.18
0.989	0.926	1	0.978	0.977	0.994	0.992	0.896	0.977	-0.03	-0.13	-0.13	-0.21	-0.21	-0.12
0.957	0.94	0.978	1	0.939	0.975	0.983	0.874	0.999	-0.01	-0.11	-0.11	-0.22	-0.19	-0.11
0.951	0.89	0.977	0.939	1	0.976	0.96	0.795	0.939	-0.02	-0.1	-0.11	-0.19	-0.19	-0.11
0.988	0.953	0.994	0.975	0.976	1	0.991	0.88	0.972	-0.02	-0.12	-0.12	-0.2	-0.2	-0.12
0.983	0.942	0.992	0.983	0.96	0.991	1	0.904	0.979	0.056	-0.05	-0.04	-0.17	-0.12	-0.04
0.927	0.83	0.896	0.874	0.795	0.88	0.904	1	0.868	-0.04	-0.13	-0.12	-0.2	-0.17	-0.11
0.954	0.935	0.977	0.999	0.939	0.972	0.979	0.868	1	-0.02	-0.12	-0.12	-0.21	-0.2	-0.11
-0.05	-0.08	-0.03	-0.01	-0.02	-0.02	0.056	-0.04	-0.02	1	0.981	0.986	0.444	0.968	0.984
-0.15	-0.2	-0.13	-0.11	-0.1	-0.12	-0.05	-0.13	-0.12	0.981	1	0.977	0.463	0.975	0.984
-0.14	-0.15	-0.13	-0.11	-0.11	-0.12	-0.04	-0.12	-0.12	0.986	0.977	1	0.486	0.987	0.987
-0.2	-0.23	-0.21	-0.22	-0.19	-0.2	-0.17	-0.2	-0.21	0.444	0.463	0.486	1	0.501	0.5
-0.21	-0.23	-0.21	-0.19	-0.19	-0.2	-0.12	-0.17	-0.2	0.968	0.975	0.987	0.501	1	0.98
-0.13	-0.18	-0.12	-0.11	-0.11	-0.12	-0.04	-0.11	-0.11	0.984	0.984	0.987	0.5	0.98	1

表8は、経常利益と純利益を出力とする12個の比データの相関行列である。経常利益と純利益を分子としているが、66個中僅か11個しか0.9以上の相関がない。この点で、経常利益と純利益の関係が少ないので6入力3出力でDEA分析を行ってもよさそうだ。

表8 経常利益と純利益を出力とする比の相関行列

行	経常 /人	経常 /店	経常/ 費用	経常/ 資金	経常/ 国債	経常/ 預金	純益 /人	純益 /店	純益/ 費用	純益/ 資金	純益/ 国債	純益/ 預金
経常/人	1	0.765	0.655	0.539	0.88	0.882	0.762	0.712	0.484	0.177	0.739	0.744
経常/店	0.765	1	0.157	0.231	0.82	0.931	0.713	0.984	0.181	0.178	0.928	0.925
経常/費用	0.655	0.157	1	0.626	0.613	0.384	0.165	0.07	0.298	-0.03	0.169	0.097
経常/資金	0.539	0.231	0.626	1	0.565	0.459	0.127	0.154	0.1	0.644	0.261	0.24
経常/国債	0.88	0.82	0.613	0.565	1	0.928	0.529	0.758	0.158	0.144	0.837	0.75
経常/預金	0.882	0.931	0.384	0.459	0.928	1	0.733	0.911	0.266	0.278	0.93	0.933
純益/人	0.762	0.713	0.165	0.127	0.529	0.733	1	0.767	0.738	0.303	0.738	0.841
純益/店	0.712	0.984	0.07	0.154	0.758	0.911	0.767	1	0.244	0.22	0.94	0.959
純益/費用	0.484	0.181	0.298	0.1	0.158	0.266	0.738	0.244	1	0.283	0.306	0.366
純益/資金	0.177	0.178	-0.03	0.644	0.144	0.278	0.303	0.22	0.283	1	0.279	0.355
純益/国債	0.739	0.928	0.169	0.261	0.837	0.93	0.738	0.94	0.306	0.279	1	0.949
純益/預金	0.744	0.925	0.097	0.24	0.75	0.933	0.841	0.959	0.366	0.355	0.949	1

4.3 Ward法によるクラスター分析

クラスター分析は理論が難しくないで、大学の初心者の統計教育に取り入れられたりしている。しかし扱う距離と手法が沢山あり、どれを用いるかの選択が重要である。距離は取りあえずユークリッド距離を用い、手法は群平均法かWard法を用いればよいが、それが絶対

正しいとは言い切れない。また、データを正規化しないと間違っただけの結果を導くことも多い。さらに、ケースのクラスター分析は一般に意味のある解釈は難しいので変数のクラスターを中心にすべきである。またJMPが提供するモザイク図は、赤から濃紺に色分けされ、赤い部分に関係するケースと変数の関係が強いと考えて解釈すればよいが、論文などではモノクロになり、読者に違いを伝えにくい。これまでケースを幾つかのクラスターに分けた後、マーカーを付加し、それを主成分分析のスコアプロットに投影して解釈することを進めてきた。しかしDEA分析を行った場合、効率的なDEAと非効率なDEAにマーカーをつけて表示することで、探索的なクラスター分析が仮説検証的なクラスター分析として扱うことができる。

4.3.1 元データの9変数と18個の比データのクラスター分析

(1) 元データの9変数と2個の効率値の分析

図7は、元データの9変数と2個の効率値に関するクラスター分析である。変数のクラスターは、2個の効率値と9個の変数に完全に分かれている。すなわち、元データには効率的な情報が含まれていない。ケースの上から3個のメガバンク3行は、元データの9変数の値が大きいことがモザイク図の赤い(黒白の濃い部分)ことから分かる。次の4(みずほ信託)、8(セブン銀行)、19(巣鴨信金)、20(城南信金)はCCRモデルで効率的であるが、手本に不適な銀行である。次の5,9,12,10,11,18,17は、効率的でない7銀行である。次の地銀4行は前に比べて良いと考えられる。次の6と7は、変数のクラスターのCCRモデルの効率値と逆CCRモデルの非効率値が高い住友信託銀行とオリックス銀行であるが、住友信託銀行は値が小さいことが分かる。9変数は、従業員数、経常費用、預金残高、貸出金と国債保有額が一つのクラスターになる。そして資本金と純利益が最初にクラスターになっているが、これらとクラスターになり、それに本支店数、経常利益が加わり一つのクラスターになっている。以上の分析は、DEA分析を行ったあとではあまり意味がないように考えられる。また今回ケースのクラスター分析は比較的解釈しやすいが、一般的にはこのように分かりやすい例はまれである。

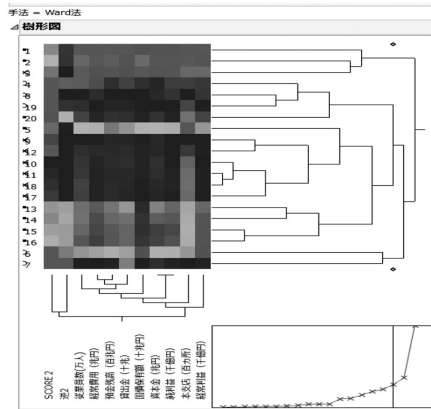


図7 元データのWard法による8変数と2個の効率値のクラスター分析

(2) 18個の比データと2個の効率値の分析

図8は18個の比データと2個の効率値のクラスター分析である。SCOREと経常利益／資本金がクラスターになりそのあと3個の比（純益／資本金，貸出金／資本金，経常利益／経常費用）で最初のクラスターを形成している。2番目は9個の比である。最後は逆SCORE（逆2）と貸出金を出力とする5個の比（貸出金／従業員数，貸出金／本支店数，貸出金／預金残高，貸出金／経常費用，貸出金／国債残高）がクラスターになっている。より詳細な実証研究が必要であるが，少なくとも元データより納得のいく結果である。ケースのクラスターでは真ん中に6の住友信託，20の城南信金，4のみずほ信託，19の巢鴨信金がSCOREと最初にクラスター化された5変数と関係が深い。最後のオリックス銀行（8）はSCOREと3番目の変数のクラスターに含まれる逆SCOREを含む6変数と関係があり，8のセブン銀行は，SCOREと2番目の9個の比と関係がある。すなわち従業員，国債，預金，本支店数が少ない割に経常利益が大きく，本支店数，国債，預金残高，従業員数が少ない割に純益が多い。また純益／経常費用はそれほど強い関係ではないという関係を比データは見事にとらえている。

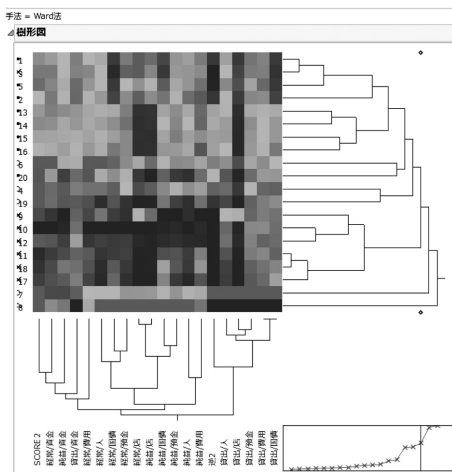


図8 比データの12変数と2個の効率値のクラスター分析

4.3.2 8個の重みと2個の効率値のクラスター分析

図9は元データのWard法による8個の重みと2個の効率値のクラスター分析である。大きく3つのクラスターが考えられる。SCOREと逆SCOREがクラスター化され、W2（本支店数）、W3（資本金）が順次クラスター化される。次にW1（従業員数）、W5（国債保有額）、W6（預金残高）のクラスターがある。最後は、W4（資本金）とW7（貸出金）、W8（経常利益）が一番クラスターとして距離が小さくまとまっている。ただし赤（濃い）の部分が散在し、ケースと変数の対応が難しい。

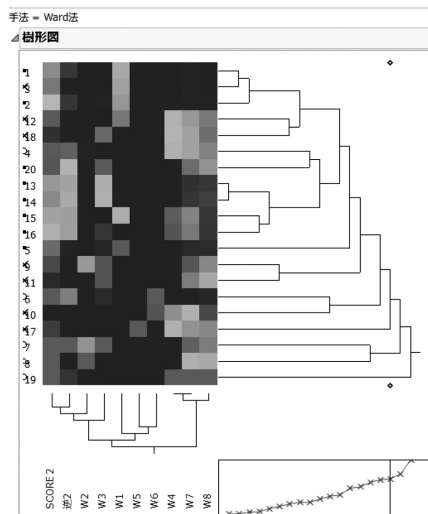


図9 元データのWard法による9個の重みと効率値のクラスター分析

4.3.3 クロス効率値と2個の効率値のクラスター分析

図10は20行のクロス効率値と2個の効率値のクラスター分析である。SCORE2と巢鴨信金がかなりの距離でクラスター化され、S4,S10,S12,S15,S16,S18,S17とクラスターになる。逆SCOREはS11,S13,S14,S20と2番目のクラスターになる。3番目はS1,S2,S5,S3がクラスターになり、最後はS7, S9, S8がクラスターになっている。

モザイク図から、みずほ信託と巢鴨信金はSCOREと関係があるが、三井住友信託とオリックス銀行はSCOREと逆SCOREの両方と関係があることが分る。真ん中にある城南銀行はSCOREと逆SCOREに、セブン銀行はSCOREと3番目のクラスターと関係が深い。

ただし、重みとクロス効率値のクラスター分析はあまり意味がないようだ。

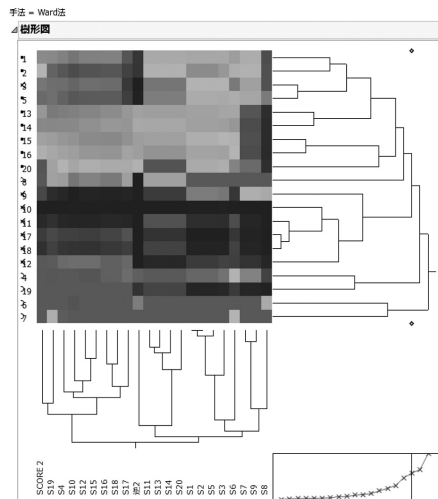


図10 元データのWard法による20行のクラスター分析

4.4 主成分分析

主成分分析は、クラスター分析のように距離や手法の選択に迷うことがない。また、因子分析のように潜在因子数の決定の違いに影響されないので、データの把握に安心して使える手法である。

4.4.1 元データと比データの主成分分析

(1) 8変数と2効率値の分析

図11は元データの主成分分析である。第1主成分の固有値が8.59で第2主成分が1.51と1以上のものは第2主成分までである。左がスコアプロット（表10の散布図）で、右が因子負荷プロットである。因子負荷量プロットは、表9に示す因子負荷量（各変数と主成分の相関係数）

をプロットしたものである。SCOREと逆SCOREは第2象限に布置し、第2主成分とほぼ0.84から0.86の高い正の相関があるが、第1主成分とは -0.12 と -0.17 である。8個の変数と第1主成分の相関は全て0.9以上と高い。第2主成分とは従業員数が0.04、経常費用が0.01、貸出金が0.09で経常利益は0.23であり1象限に布置する。本支店は -0.04 、資本金は -0.002 、国債保有は -0.06 で第4象限に布置する。この違いは、第2軸の解釈に利用すればよい。

この結果は、左のスコアプロットの解釈に利用できる。メガバンク3行は全ての変数値が大きく、第1次軸の正に布置し、負の多くの銀行と対立軸で考えることができる。第2軸は6の住友信託と7のオリックスが経常利益が良く、負の銀行と対比して考えることができる。すなわち、各象限でまとまった銀行がデータとして似た傾向を示す。すなわち表10の2象限にあるオリックスを除く7行は経常利益は評価できるが規模の小さい銀行である。第4象限の5の三菱UFJ信託と14の千葉銀行と残りの6行は経常利益が他の7変数に比べて少ないと考えられる。

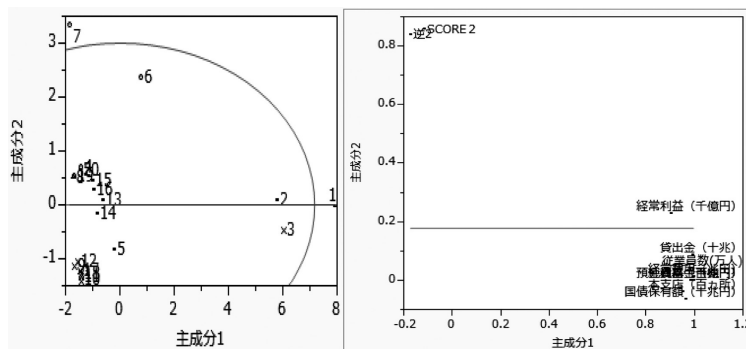


図11 元データの主成分分析 (左:スコアプロット, 右:因子負荷量)

表9 因子負荷量

行	主成分1	主成分2	象限
SCORE 2	-0.117	0.858	2
逆2	-0.170	0.839	2
従業員数 (万人)	0.992	0.041	1
本支店 (百カ所)	0.954	-0.041	4
経常費用 (兆円)	0.996	0.014	1
資本金 (兆円)	0.983	-0.002	4
国債保有額 (十兆円)	0.964	-0.064	4
預金残高 (百兆円)	0.997	0.000	1
貸出金 (十兆)	0.993	0.085	1
経常利益 (千億円)	0.905	0.228	1

表10 主成分のスコア

SN	銀行	主成分1	主成分2	象限
1	三菱東京UFJ銀行	7.574	-0.034	4
2	三井住友銀行	5.253	0.080	1
3	みずほ銀行	5.742	-0.470	4
4	みずほ信託銀行	-1.339	0.684	2
5	三菱UFJ信託銀行	-0.245	-0.828	3
6	三井住友信託銀行	0.848	2.363	1
7	オリックス銀行	-1.777	3.328	2
8	セブン銀行	-1.637	0.500	2
9	ソニー銀行	-1.561	-1.143	3
10	東京都民銀行	-1.328	-1.410	3
11	八千代銀行	-1.326	-1.345	3
12	東京スター銀行	-1.447	-1.046	3
13	横浜銀行	-0.581	0.072	2
14	千葉銀行	-0.752	-0.163	3
15	福岡銀行	-0.919	0.438	2
16	静岡銀行	-0.910	0.286	2
17	多摩信用金庫	-1.347	-1.213	3
18	城北信用金庫	-1.322	-1.288	3
19	巣鴨信用金庫	-1.582	0.536	2
20	城南信用金庫	-1.345	0.653	2

(2) 12個の比の主成分分析

図12は12個の比率の主成分分析である。固有値が1以上のものは3個ある。表11の散布図である図の右の因子負荷プロットから、SCOREと逆SCOREと経常利益を出力とする6個の比は、1象限にある。貸出金を出力とする6個の比は、4象限にある。

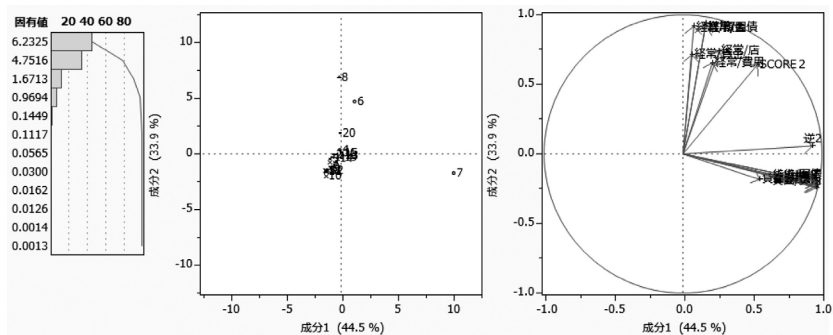


図12 12個の比データの主成分分析 (左:スコアプロット, 右:因子負荷量)

表11 比データの因子負荷量

行	主成分1	主成分2	象限
SCORE 2	0.534	0.641	1
逆2	0.930	0.058	1
貸出／人	0.952	-0.232	4
貸出／店	0.954	-0.198	4
貸出／費用	0.962	-0.241	4
貸出／資金	0.550	-0.183	4
貸出／国債	0.967	-0.199	4
貸出／預金	0.964	-0.222	4
経常／人	0.162	0.922	1
経常／店	0.259	0.739	1
経常／費用	0.212	0.656	1
経常／資金	0.065	0.712	1
経常／国債	0.164	0.943	1
経常／預金	0.077	0.917	1

これは表12の主成分スコアとその散布図のスコアプロットの解釈に利用できる。1象限に住友信託銀行と福岡銀行と静岡銀行の3行があるが、地銀2行はほぼ原点に近い。これらは経常利益の比が高い。4象限にオリックス銀行と5行があるが、貸出比率が高いことを表す。単純化すれば、3象限にある10行は経常利益の比が悪く、2象限のセブン銀行と城南信金の2行は貸出比率が悪いことは良く理解できる。すなわちDEAの効率値は、経常利益の比に重きをおいて、主成分分析では1象限と3象限の違いで表現していると考えられる。

表12 比のスコアプロット

SN	銀行	主成分1	主成分2	象限
1	三菱東京UFJ銀行	-0.872	-0.329	3
2	三井住友銀行	-0.689	-0.106	3
3	みずほ銀行	-0.963	-0.747	3
4	みずほ信託銀行	-0.068	0.293	4
5	三菱UFJ信託銀行	-1.008	-0.477	3
6	三井住友信託銀行	1.218	4.655	1
7	オリックス銀行	10.194	-1.738	4
8	セブン銀行	-0.187	6.795	2
9	ソニー銀行	-0.906	-1.280	3
10	東京都民銀行	-1.357	-2.008	3
11	八千代銀行	-1.362	-1.568	3
12	東京スター銀行	-1.186	-1.473	4
13	横浜銀行	0.189	-0.139	4
14	千葉銀行	-0.328	-0.436	3
15	福岡銀行	0.042	0.068	1
16	静岡銀行	0.031	0.052	1
17	多摩信用金庫	-1.345	-1.501	3
18	城北信用金庫	-1.384	-1.501	3
19	巣鴨信用金庫	0.077	-0.369	4
20	城南信用金庫	-0.098	1.810	2

4.4.2 重みの分析

図13は重みの主成分分析である。第1主成分から第4主成分までの固有値が2.92, 2.14, 1.27, 1.15で1以上である。真ん中がスコアプロット（表14の主成分スコアの散布図）で、右が因子負荷プロットである。因子負荷量プロットは、表13に示す因子負荷量をプロットしたものである。変数値と異なり、重みなのでSCOREと逆SCOREとW2,W3,W8が第1象限、W6が第2象限、W1が第3象限、W4,W5,W7が第4象限に布置している。左のスコアプロット上の20行もほぼこれに対応していると考えてよいであろう。

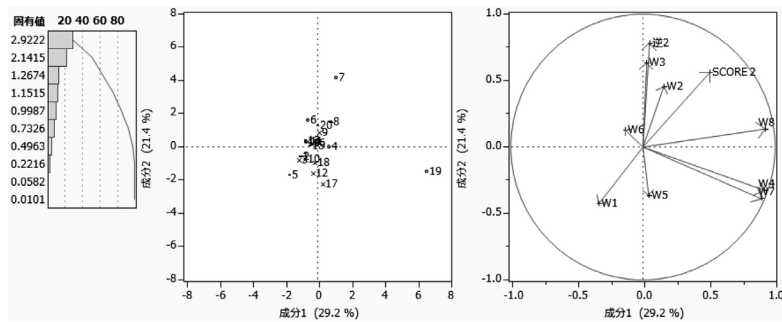


図13 比データの主成分分析（左：スコアプロット，右：因子負荷プロット）

表13の因子負荷量からSCOREと逆SCOREと本支店数，経常費用，経常利益が第1象限にある。表14からみずほ信託とインターネット銀行3行が1象限に布置される。従業員が第3象限に布置し，メガバンク3行のほかに4行が布置されるが解釈に苦しむ。

以上から，重みの解釈は比データよりはっきりしないようだ。

表13 重みの因子負荷量

行	主成分1	主成分2	象限
SCORE 2	0.876	-0.232	4
逆2	0.786	0.065	1
S1	0.837	0.515	1
S2	0.837	0.515	1
S3	0.849	0.495	1
S4	0.788	-0.592	4
S5	0.837	0.515	1
S6	0.708	0.388	1
S7	0.797	0.456	1
S8	0.604	0.470	1
S9	0.797	0.456	1
S10	0.792	-0.590	4
S11	0.847	0.209	1
S12	0.823	-0.556	4

S13	0.847	0.209	1
S14	0.847	0.209	1
S15	0.823	-0.556	4
S16	0.850	-0.518	4
S17	0.856	-0.489	4
S18	0.850	-0.518	4
S19	0.709	-0.577	4
S20	0.847	0.209	1

表14 主成分スコア

SN	銀行	主成分1	主成分2	象限
1	三菱東京UFJ銀行	-1.056	-0.597	3
2	三井住友銀行	-1.030	-0.592	3
3	みずほ銀行	-1.141	-0.774	3
4	みずほ信託銀行	0.716	0.009	1
5	三菱UFJ信託銀行	-1.727	-1.697	3
6	三井住友信託銀行	-0.592	1.621	2
7	オリックス銀行	1.128	4.177	1
8	セブン銀行	0.821	1.500	1
9	ソニー銀行	0.130	0.852	1
10	東京都民銀行	-0.767	-0.772	3
11	八千代銀行	-0.600	0.300	2
12	東京スター銀行	-0.253	-1.614	3
13	横浜銀行	-0.740	0.389	2
14	千葉銀行	-0.750	0.293	2
15	福岡銀行	-0.473	0.067	2
16	静岡銀行	-0.430	0.225	2
17	多摩信用金庫	0.316	-2.262	4
18	城北信用金庫	-0.151	-0.942	3
19	巢鴨信用金庫	6.602	-1.479	4
20	城南信用金庫	-0.003	1.296	2

4.4.3 クロス効率値の主成分分析

図14はクロス効率値の主成分分析である。第1主成分から第3主成分までの固有値が14.5, 4.49, 1.65で1以上である。右の因子負荷プロットと表15から、逆SCOREとS1からS3,S5からS9, S11,S13,S14,S20が第1象限、SCOREとS4,S10,S12,S15からS19が第4象限に布置している。これに対してスコアプロットでは、6から7の三井住友信託、オリックス、セブン銀行、20の城南信金が布置している。第3象限にメガバンク3行、5の三菱UFJ信託、12の東京スター銀行、17と18の多摩信金と城北信金が布置している。因子負荷量とスコアプロットは対応する部分もあるが、比データほど必ずしも対応していない。

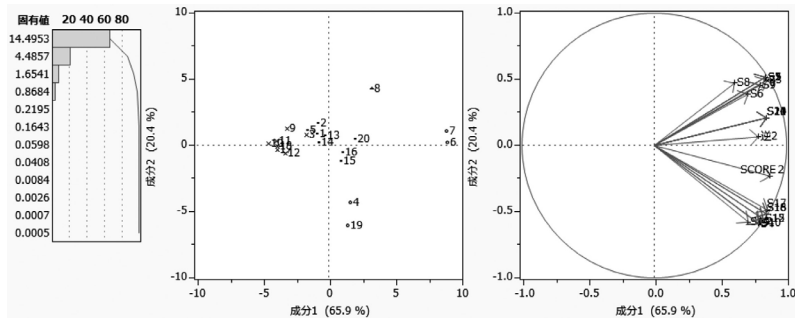


図14 クロス効率値の主成分分析（左：スコアプロット，右：因子負荷プロット）

表15 因子負荷量

行	主成分1	主成分2	象限
SCORE 2	0.876	-0.232	4
逆2	0.786	0.065	1
S1	0.837	0.515	1
S2	0.837	0.515	1
S3	0.849	0.495	1
S4	0.788	-0.592	4
S5	0.837	0.515	1
S6	0.708	0.388	1
S7	0.797	0.456	1
S8	0.604	0.470	1
S9	0.797	0.456	1
S10	0.792	-0.590	4
S11	0.847	0.209	1
S12	0.823	-0.556	4
S13	0.847	0.209	1
S14	0.847	0.209	1
S15	0.823	-0.556	4
S16	0.850	-0.518	4
S17	0.856	-0.489	4
S18	0.850	-0.518	4
S19	0.709	-0.577	4
S20	0.847	0.209	1

表16 主成分スコア

SN	銀行	主成分1	主成分2	象限
1	三菱東京UFJ銀行	-0.846	0.845	3
2	三井住友銀行	-0.800	1.698	3
3	みずほ銀行	-1.654	0.780	3
4	みずほ信託銀行	1.685	-4.354	4
5	三菱UFJ信託銀行	-1.577	1.122	3
6	三井住友信託銀行	9.014	0.175	1
7	オリックス銀行	8.909	1.063	1
8	セブン銀行	3.292	4.247	1
9	ソニー銀行	-3.114	1.286	2
10	東京都民銀行	-4.534	0.152	2
11	八千代銀行	-3.910	0.333	2
12	東京スター銀行	-3.235	-0.568	3
13	横浜銀行	-0.263	0.766	2
14	千葉銀行	-0.728	0.167	2
15	福岡銀行	0.951	-1.228	4
16	静岡銀行	1.044	-0.524	4
17	多摩信用金庫	-3.843	-0.335	3
18	城北信用金庫	-3.868	-0.061	3
19	巣鴨信用金庫	1.466	-6.061	4
20	城南信用金庫	2.011	0.494	1

4.5 重回帰分析

3個の出力とSCOREと逆SCOREの5個を目的変数として、6個の元データの入力変数と18個の比データを説明変数とした重回帰分析を比較検討する。

4.5.1 貸出金の重回帰分析

図15は貸出金を6変数で重回帰分析した結果である。決定係数は0.99と申し分ない。回帰係数のp値は0.15以上で、VIFは31.7から403.6の間である。VIFは多重共線性を表す指標の一つであり、1個の説明変数を残りの説明変数で回帰した際の決定係数を R^2_i とすれば $VIF_i = 1 / (1 - R^2_i)$ で表される。例えば預金残高の $VIF = 1 / (1 - R^2_i) = 403.6$ だから、 $R^2_i = 0.997522$ になる。すなわち、預金残高は他の説明変数のほぼ線形式で表され、回帰係数の標準誤差が大きくなり（本事例では小さい）95%信頼区間は0を含んでいる。目安はないが100以上程度が望ましい。プロット図から、オリックスだけが外れ値になっている。

すなわち、メガバンク、信託銀行、地方銀行、信託銀行の順に入力変数のスケールで貸出金の予測がすみ分けされているのが金融業で、当然と言えば当然の結果である。

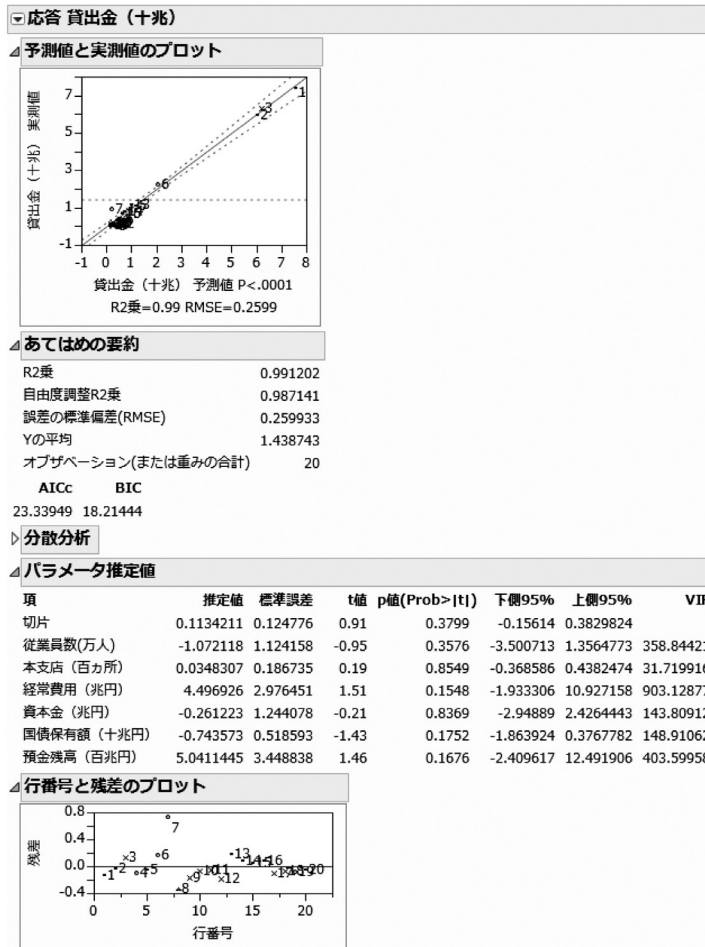


図 15 貸出金を6変数で重回帰分析

残差のプロットではオリックス銀行や三井住友信託が正の残差で、負の残差で×印の銀行が目立つようだ。これは表16の残差で詳細を論じる。一番上の予測値と残差のプロットから、メガバンク3行の残差は小さいが、予測値と実測値が大きい。住友信託とオリックスはわずかながら予測値の95%信頼区間の正の方に外れ値になっている。金融業においては、それほど大きな残差は現れないようだ。

図16は、説明変数の全ての組み合わせ ($=2^6 - 1$) の中から、AICとBICが最小の本支店数と経常費用の2変数を選んだ。多重共線性はこのように変数選択を行えば一般的には解消されVIFも7と小さくなる。少なくとも50以下で多重共線性ありと判定することは間違っている。決定係数は0.988で悪くなく、標準誤差は小さくt検定は5%で棄却されているが、プロ

ット図に図15と大きな違いはない。

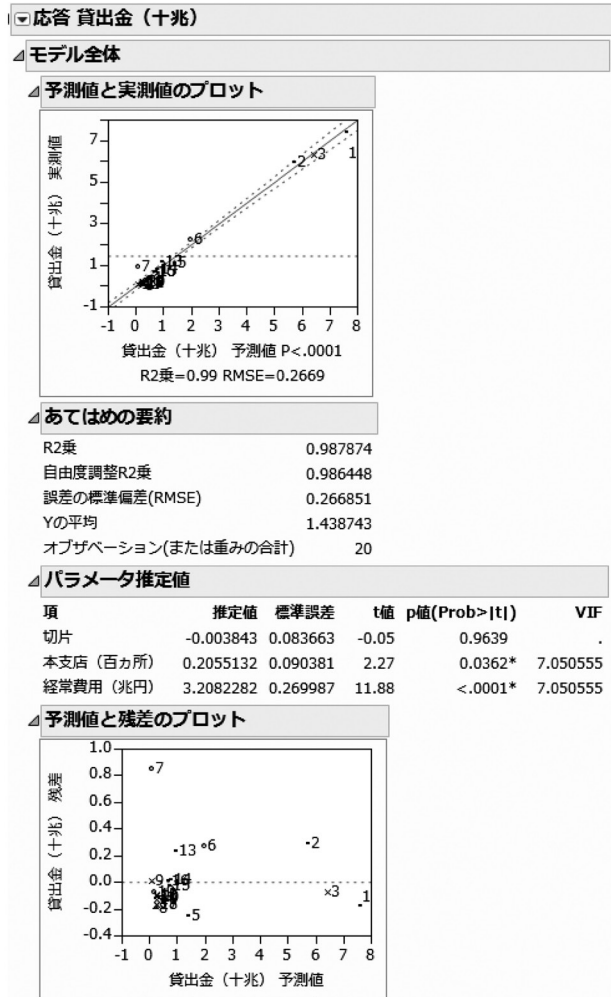


図16 本支店数と経常費用の2変数モデル

4.5.2 経常利益の重回帰分析

図17は経常利益を6変数で重回帰分析した結果である。決定係数は0.98と申し分ない。回帰係数のp値は経常費用と国債が棄却される。VIFは32から404の間である。

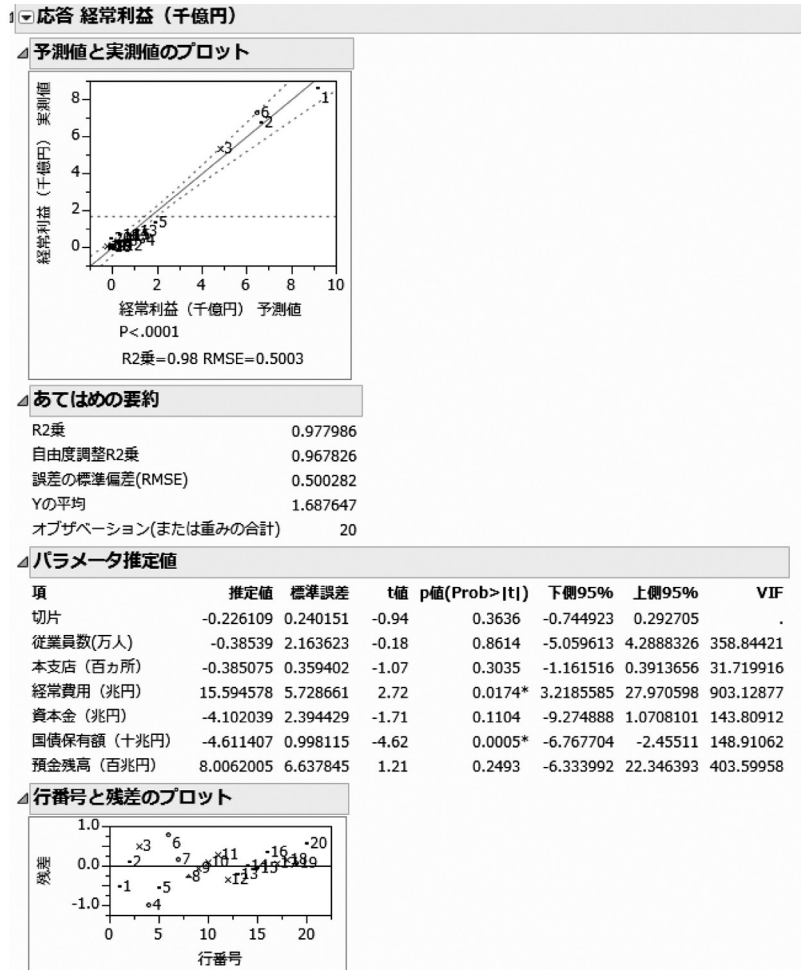


図17 経常利益を6変数で重回帰分析した結果

図18は、説明変数の全ての組み合わせ ($= 2^6 - 1$) の中から、AICとBICが最小の3変数を選んだ。多重共線性は変数選択を行っても25から69もあり、経常費用の69は多重共線性が疑われるボーダーであろう。P値は3変数とも棄却されている。三井住友信託は、経常費用に関してメガバンク3行のレベルにあることが分かる。予測値と残差のプロットは、大きく差が見て取れるが、効率的なものが正で、非効率的なものが負という明確な傾向が認められない。

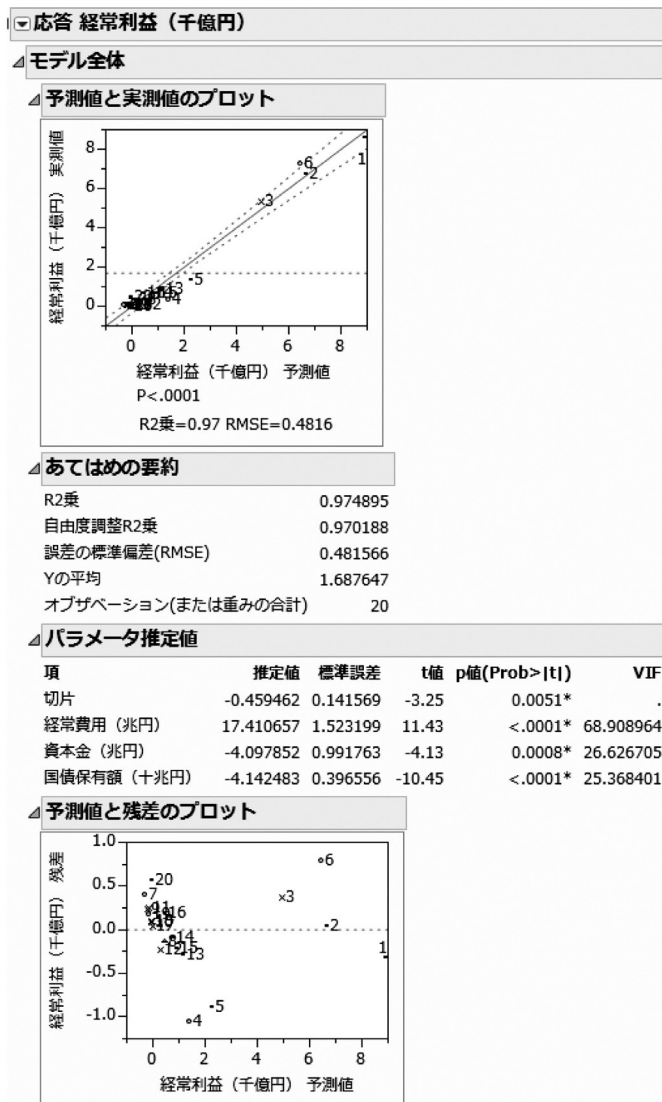


図18 3変数モデル

4.5.3 純利益の重回帰分析

図19は純利益を6変数で重回帰分析した結果である。決定係数は0.998と申し分ない。回帰係数のp値は資本金だけが棄却される。VIFは32から404の間である。全てのモデルの選択でもこの6変数が選ばれた。すなわち純利益を予測するのに6個の説明変数がそれぞれ必要であり、この点で純利益で経営効率性の違いにアプローチすることが適しているのかもしれない。

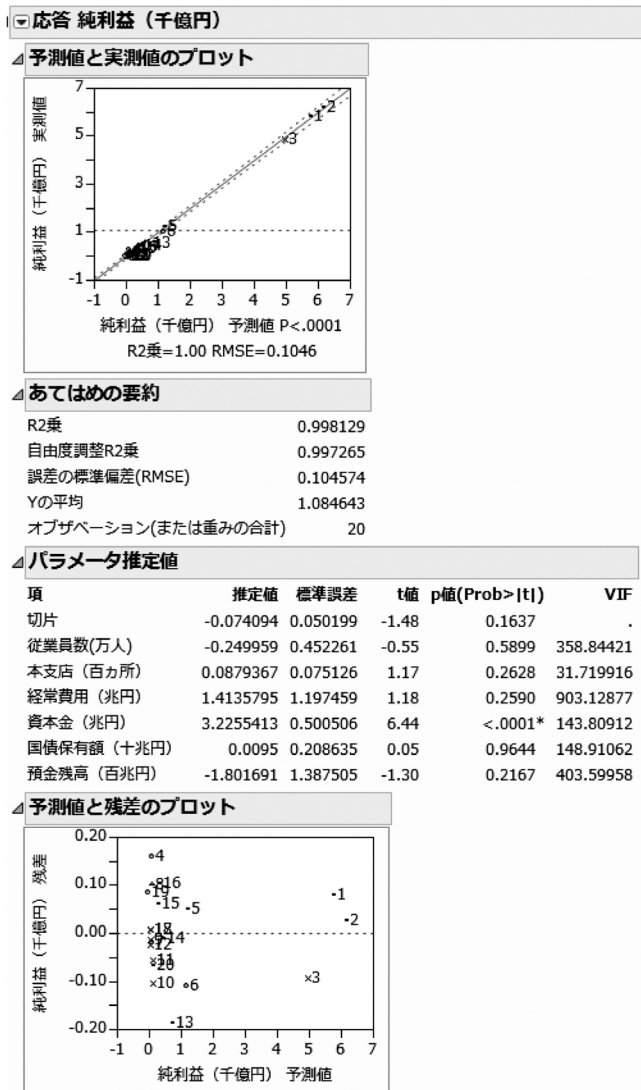


図19 純利益を6変数で重回帰分析した結果

4.5.4 SCOREの重回帰分析

図20はSCOREを6変数で重回帰分析した結果である。決定係数は0.29と悪い。回帰係数のp値は全て棄却されない。またVIFも32から903と幅広い。予測値と実測値のプロットは、信頼区間から正の方に外れるのは7, 19, 20で、負の方に外れるのは9, 10, 12とDEAで効率的なものと非効率なものが選ばれている。すなわち、DEAの評価基準は6個の入力変数での予測がうまくいかない。



図20 SCOREを6変数で重回帰分析した結果

変数選択で図21の本支店数だけが選ばれたが単回帰でも予測精度が悪い。また実測値のSCOREが0.2から1の範囲でばらついているが、予測値は0.4から0.6の間である。すなわち、CCRモデルで効率値を求める分析は、単純に元データから情報を引き出しているとは考えられない。

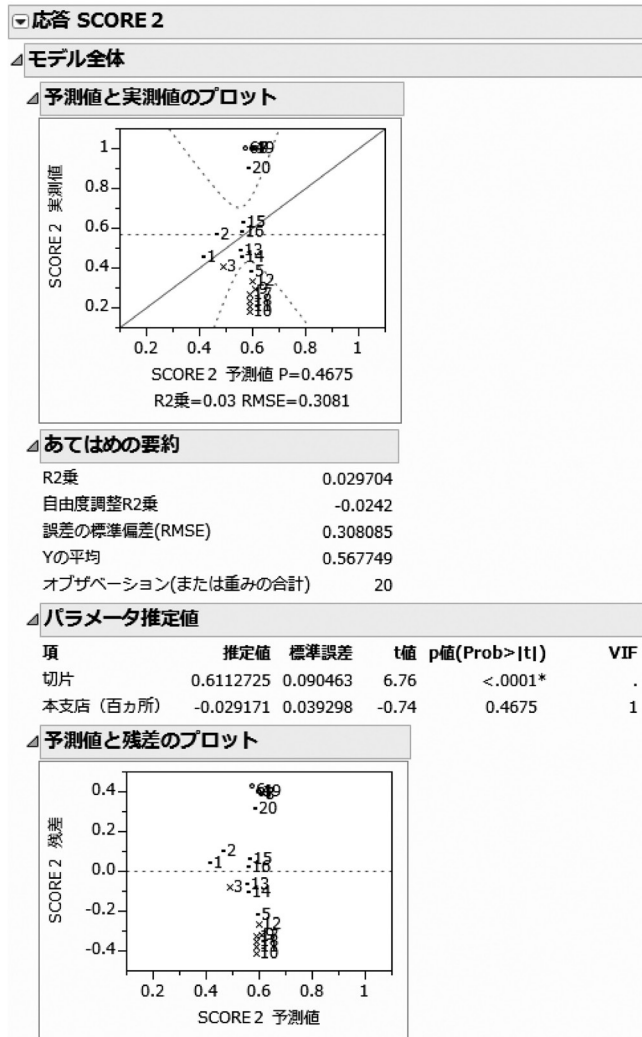


図21 変数選択で本支店数だけが選ばれたが単回帰

4.5.5 逆SCOREの重回帰分析

図22は逆SCOREを6変数で重回帰分析した結果である。決定係数は0.09と非常に悪い。回帰係数のp値は全て棄却されない。予測値と実測値のプロットは、信頼区間から外れるはSCOREに比べて少ない。すなわち、DEAの評価基準は元データで予測できない。

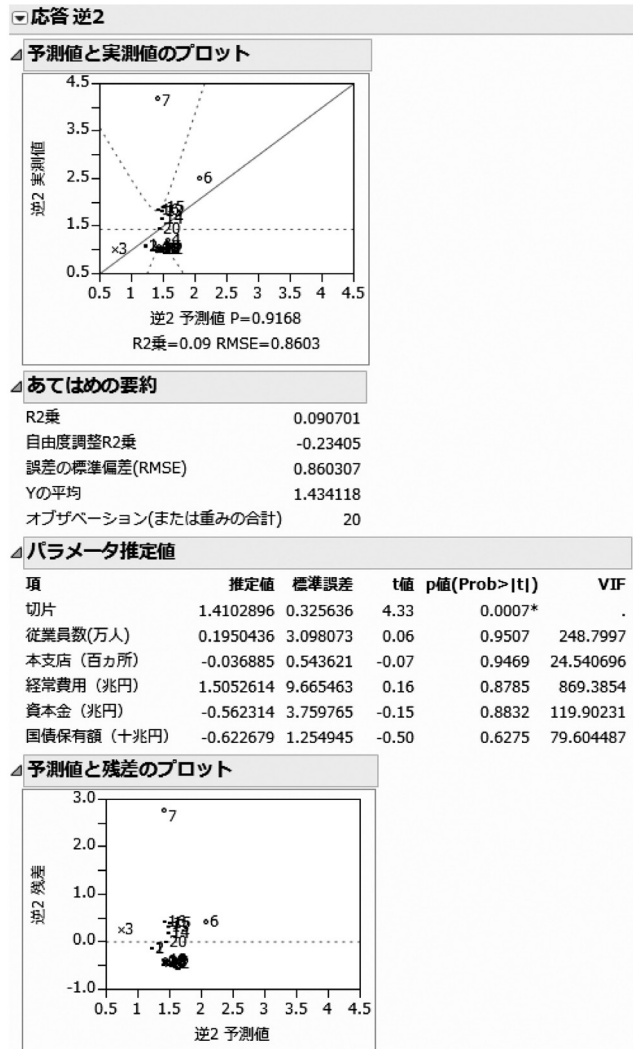


図22 逆SCOREを6変数で重回帰分析

4.5.6 SCOREと逆SCOREの比尺度による重回帰分析

図23はSCOREを18個の比変数で重回帰分析して、変数選択後の5変数の結果である。決定係数は0.98と良い。回帰係数のp値は全て棄却されVIFも2.5以下になっている。しかし予想に反して、効率的なものと非効率なものが大きな残差になっていない。

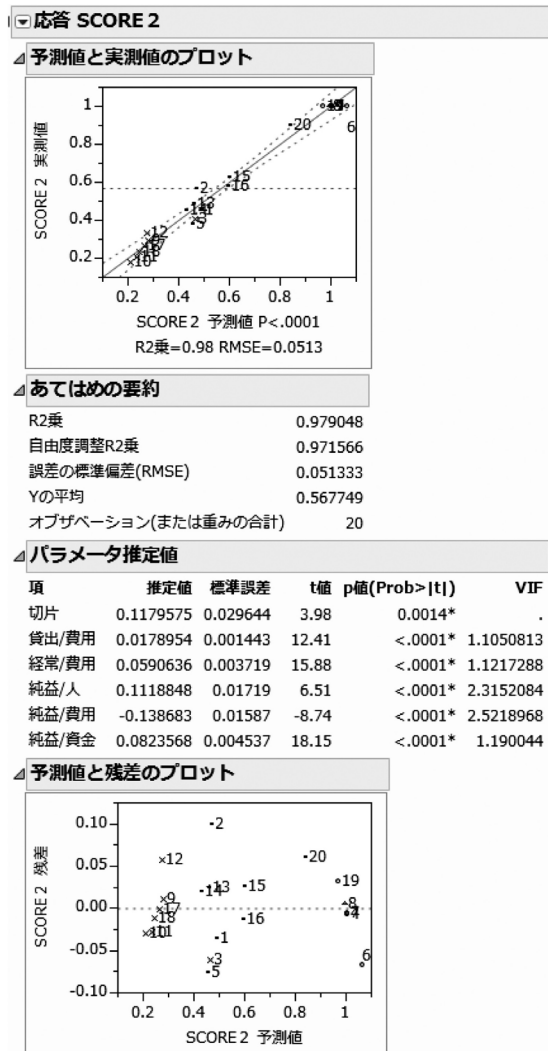


図23 SCOREを18個の比変数で重回帰分析して、変数選択後の6変数の結果

図24は逆SCOREを18個の比変数で重回帰分析して、変数選択後の5変数の結果である。決定係数は0.98と良い。回帰係数のp値は全て棄却されているが、VIFは4.3から84である。しかし、オリックスは予測誤差がほぼ0で、住友信託と巣鴨信金は正の残差であるが、非効率な多摩信金も正の残差である。

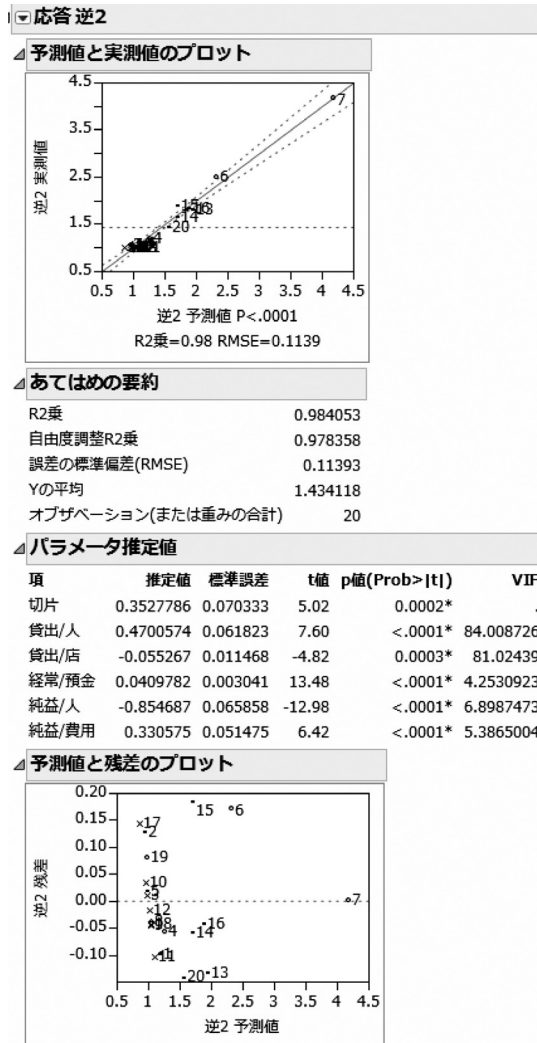


図24 逆SCOREを36個の比変数で重回帰分析して、変数選択後の5変数の結果

5. まとめ

本研究では、DEA分析の結果を統計分析で検証することで、DEAと統計分析の両方に大きな成果をもたらす可能性を示した。単に元データのFだけで統計分析しても、分析の解釈は手探り状態で不安である。それが、DEAで分かった効率的と非効率的なDMUで層別して、統計手法で検証することで豊富な統計手法でDEAの結果を解釈できる。さらに、F以外に重みW、クロス効率値S、SCOREと入力と出力の比データといった分析を行うことができる。

今回分かったことは、DEA効率値は、Fから説明できず比データを用いて初めて説明できるようだ。しかし、重みやクロス効率値から有用な情報は分からなかった。CCRモデルで重みWを求めることは、Fから総合化された特殊な比情報を抽出していることも考えられる。

DEA手法は、多くの学生にとって問題点を教えてくれる。この点で、就活の一環としての企業研究を卒業研究のテーマとして捉えることも必要であろう。

(成蹊大学経済学部教授)

参考文献

- [1] 新村秀一 (2004). 『JMP活用 統計学とおき勉強法』. 講談社.
- [2] 新村秀一 (2011). 『数理計画法による問題解決法』. 日科議連出版社.
- [3] 新村秀一 (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.
- [4] 新村秀一 (2013). DEA利用のための実践的な解説書—1986年と2011年の東京都23区の公立図書館の比較評価—. 成蹊大学経済学部論集, 44/1, 15-41.
- [5] 新村秀一 (2013). 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.
- [6] 杉山学 (2010). 『経営効率分析のためのDEAとInverted DEA—基本概念と方法論から、主観的な判断を加味できる応用モデルまで—』. 静岡学術出版.
- [7] 山田善治靖, 松井知己, 杉山学 (1994). DEAモデルに基づく新たな経営効率性分析法の提案. *Journal of the Operations Research Society of Japan*, 7,158-168.
- [8] L. Schrage (2003). "Optimizer Modeling with LINGO". LINDO Systems Inc.
- [9] S. Shinmura (2012). Relationship between the DEA cluster and the lower limit of weights. *Proceedings of DEA Symposium 2012*, pp.63-68.
- [10] K. Tone (1988). Introduction to Efficiency Analysis of a company-DEA (1). *Operations Research*, 32/12.
- [11] T. Ueda (2007). Application of Multivariate analysis for DEA. *Proceedings of DEA Symposium 2007*, pp.96-101.