

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－

新 村 秀 一

1. はじめに

本研究では、日本車44車種の販売台数（台数）と販売価格（価格）と4個の性能のデータを用いて、最初にDEA法で分析し統計手法と比較評価を行う。DEA研究では、重回帰分析はデータ全体の平均的な特徴しか捉えられないのに、DEAでは個別の特徴を評価できる点を優位点に挙げている。さらに多入力多出力のデータは正準相関分析で分析できるが、重回帰分析では台数と価格を目的変数として各性能を説明変数とした異なった重回帰分析で分析する必要がある。しかし、重回帰分析に限らず統計手法はDEAに比べてデータの制約が少なく、変数選択法の手法がある点で優れている。これらの問題点を、日本車44車種の実データで検証する。

次に、DEAの応用分野として価格性能比の新しい方法論を検討する。DEAを多入力と多出力の総合特性値の比を評価する手法として新しく位置付ける。この応用として、複数の性能を総合化して、価格性能比を評価する。価格を入力とし総合化された性能を出力とするCCRモデルは、価格の割に性能のよい車をユーザーが購入できると考えるとユーザー側の視点に立った車の評価ができる。これに対して逆CCRモデルは、総合化された性能を入力とし価格を出力とするCCRモデルと等価である。この場合、性能の割に価格が高い車種が効率的であり、企業側に立った評価であり、ブランド力等と関係しているのではないかと考えられる。

本研究の応用として、自動車よりも携帯電話やPCなどの価格の安い多くの日用品の価格性能比の検討に適用する方が、結果が良いと考える。自動車はここで取り上げた項目では捉えきれない製品であることは、各社のカタログを比較するだけでも数多くの異なった項目があることから分かる。また車の専門雑誌が出ていること自体、他の日用製品と異なっていることが分かる。そのような製品を対象にして部外者が分析しても、得られた結果を実際に利用することは難しい。これに対して、同じ企業で多数の製品が頻繁にモデルチェンジされている携帯電話やPC等は、カタログに取り上げられている仕様はほぼ同じで、ファンのための専門雑誌も少ないので、本研究で提案する方法で評価を行うことは有益と考える。

今回自動車に全く不案内で知識のない筆者が、院生の集めた特定月のデータを用いて自動車に十分な知識と見識ないままDEAと統計手法で分析を行った。データに基づいた分析手法の方法論に関しての結果の責任は筆者にあるが、それほど間違った結果でない自信がある。問題点があれば、是非指摘いただきたい。

2. 分析データ

2011年度の成蹊大学経済経営学研究科の院生の岡野いつ佳さんは、「データ分析 (2単位)」¹で日本車44車種のデータを集め統計レポートを提出してくれた。2.1が彼女の課題レポートの3章の集めた変数に関する記述である (一部修正)。ここでは、このデータを筆者が統計分析しなおして、その結果とDEA法の分析結果を対比して、両手法を比較検討する。

ただし、2011年度の月別の順位をみると入れ替わりがあるので、本来は年間の順位であることが望ましい。また、総排気量 (排気)、燃費、CO₂排出量 (CO₂) は、カタログにも複数の値があり、各メーカーのHPから非専門家の一つに決めるのは問題が出てくることも考えられる。さらに、普通車で12位のマツダのデミオを省いて29車種にしていることもマツダは1車種しかないのでは理解できるが問題である。また自動車のような高額商品で技術要素の多い製品では、本研究で取り上げた項目だけの分析では専門家の要求水準から程遠く、現実の一面をとらえた入門の分析と考えている。以上は「データ分析」という2単位科目の提出レポートのテーマとして自動車を取り上げると聞いて、事前に十分な指導をできなかった自戒と反省点である。またある水準以上のレポートであれば、2003年以前は筆者のHPに掲載して学生の努力を顕彰してきたが、2005年以降それができていないことの反省でもある。今後は、頑張った学生の成果はそのままHP等で紹介するか、分析しなおしたものを学生にフィードバックをできるだけ速やかに行いたい。昨年度は、卒業演習の黒岩さんは計算機統計学会と、上級演習の江頭さんはOR学会の評価の研究部会の学生発表会で黒岩さんと発表させることを試みた。学生の負担も大きいので今後は慎重に行う予定である。

2.1. データの基本情報

分析対象のデータは2011年12月の新車販売台数をもとに、車種のスペック情報を詳しく分析し、現在の消費者に求められているスペックを明らかにしていく。自動車販売台数のデータは、国内普通自動車は (社) 日本自動車販売協会連合会 (<http://www.jada.or.jp/index.html>) から2011年12月新車乗用車販売台数ランキング30車種のデータを使用した。また小型車はこのデータには反映されていないため、(社) 全国小型車協会連合会 (<http://www.zenkeijikyoo.or.jp/index.html>) から、2011年12月軽四輪車新車販売台数のデータを使用した。また、それぞれの車種の詳しいスペック情報は、各メーカーHPからデータを使用した。

これらのデータから販売台数上位44車種のデータを使用する。変数としては通称名、メーカー、12月販売台数、価格、定員、総排気量 (L)、燃費 (L/km)、CO₂排出量、普通車と小

¹ 同じ受講生の井上正澄さんのレポートの一部は、STN (SAS Technical News) の2012年夏号で紹介した [18] .

型車の区別, エコカー減税の有無, 駆動方式, ミッション, 燃料の計13項目を使用した。変数の詳しい説明を以下で行う。

通称名: 車に通常つけられている名前を使用した。同じ名称で異なるスペックの車種では, 燃費の良い車種のデータを使用する。

メーカー: 44車種ではメーカーはトヨタ自動車(株), 日産自動車(株), 本田技研工業(株), スズキ(株), 富士重工業(株), ダイハツ工業(株), 三菱自動車工業(株)の計7社のデータを使用した。これらの自動車会社の中にはグループ会社や提携関係もあるが, メーカーとして独立しているため, これら7社に分類した。

12月販売台数: 普通自動車から29車種, 小型車から15車種の計44車種の新車販売台数。

価格: 円で表している。オプション費用は除いて, 基本的な価格のデータを使用する。

定員: 自動車の定員数 (人)。

総排気量: 自動車の排気量 (L)。

燃費: 通常の走行で1kmにガソリンが何L必要かを表したもの。一般化するため通常走行を基本としている。走行方法が変更になるものでは, もっとも通常走行に近い走行方法のデータを使用する。

CO2排出量: 通常走行時のCO2排出量 (L) を表している。

普通車/小型車: 普通車を1, 小型車を0のダミー変数で表わす。

燃料: ガソリンの種類についてはレギュラー, プレミアム (ハイオク), 軽油の3つに分類した。データは, R, P, Lと表わす。

エコカー減税: エコカー減税の対象車の有無をダミー変数で表した。割引税率の差は記載していない。

駆動方式: 前輪駆動方式の2WDと前駆動方式の4WDとで分類した。2WDはエンジンの搭載箇所によってFF, FR, MR, RR等の呼び名もあるが, ここでは2WDと4WDの2つの区別とした。データは2と4で表わす。

ミッション: 変速装置のことを言う。CVTは無段変速機と呼ばれ, 従来の変速時に現れる衝撃がないものである。5MTはマニュアルトランスミッションとよばれ, 運転者がギアを変化させる方式のこと。4ATは自動ギア変速機能と呼ばれCVTに比べて4段階しか変速できない。

2.2. 分析データ

表2.1は, 今回用いるデータである。統計分析は統計ソフトのJMP [1] [6], DEA法はLINGO [2, 7] で作成した汎用DEAプログラム [9] を用いた。

表2.1 分析データ

通称名	メーカー	価格	台数	排気	燃費	CO2	定員	普通／ 小型	ミッシ ョン	燃料	駆動	エコ 減税
スイフト	スズキ	1,244,250	2,628	1.242	25	93	5	1	CVT	R	2	1
ソリオ	スズキ	1,477,350	2,778	1.242	24.5	95	5	1	CVT	R	2	1
インプレッサ	スバル	1,711,500	2,496	1.599	17.6	132	5	1	CVT	R	2	1
レガシィ	スバル	2,362,500	1,742	2.457	14	166	5	1	CVT	R	4	1
アルファード	トヨタ	4,428,350	2,535	2.362	17	137	7	1	CVT	R	4	1
ウィッシュ	トヨタ	2,090,000	2,493	1.797	16.4	142	7	1	CVT	R	2	1
ヴィッツ	トヨタ	1,350,000	7,750	1.329	21.8	106	5	1	CVT	R	2	1
ヴォクシー	トヨタ	2,330,000	4,678	1.986	13.6	171	7	1	CVT	R	2	1
ヴォルファイア	トヨタ	3,950,000	3,762	2.362	17	137	7	1	CVT	R	4	1
エスティマ	トヨタ	2,960,000	3,870	2.362	12.4	187	7	1	CVT	R	2	1
カムリ	トヨタ	3,040,000	1,499	2.493	23.4	99	5	1	CVT	R	2	1
カローラ	トヨタ	2,440,000	4,771	1.496	20	116	5	1	CVT	R	2	1
クラウン	トヨタ	5,400,000	3,190	3.456	14	166	5	1	CVT	P	2	1
シエンタ	トヨタ	1,655,000	2,418	1.496	17.2	135	7	1	CVT	R	2	1
ノア	トヨタ	2,520,000	2,877	1.986	13.6	171	7	1	CVT	R	2	1
パッソ	トヨタ	1,000,000	4,098	0.996	20.8	112	5	1	CVT	R	2	1
プリウス	トヨタ	2,170,000	22,998	1.797	32.6	71	5	1	CVT	R	2	1
ラクティス	トヨタ	1,440,000	3,277	1.329	18.4	126	5	1	CVT	R	2	1
エクストレイル	日産	2,607,150	1,828	1.995	14.2	185	5	1	6MT	L	4	1
エルグランド	日産	3,612,000	1,517	2.488	10.8	215	7	1	CVT	R	2	1
キューブ	日産	1,529,850	2,969	1.498	18	129	5	1	CVT	R	2	1
ジューク	日産	1,620,150	2,642	1.498	17.2	135	5	1	CVT	R	2	1
セレナ	日産	2,499,000	6,512	1.997	14.6	159	8	1	CVT	R	2	1
ティーダ	日産	1,596,000	2,169	1.498	18	116	5	1	CVT	R	2	1
ノート	日産	1,339,800	3,656	1.498	18	129	5	1	CVT	R	2	1
マーチ	日産	1,036,350	1,727	1.198	22.6	103	5	1	CVT	R	2	1
ステップワゴン	ホンダ	2,088,000	4,396	1.997	13.4	173.3	8	1	CVT	R	2	1
フィット	ホンダ	1,590,000	15,248	1.339	26	89.3	5	1	CVT	R	2	1
フリード	ホンダ	2,295,000	7,817	1.496	21.6	107.5	6	1	CVT	R	2	1
MRワゴン	スズキ	1,257,900	1,868	0.658	27	86	4	0	CVT	R	2	1
アルト	スズキ	1,029,000	6,930	0.658	32	73	4	0	CVT	R	2	1
エブリイワゴン	スズキ	1,344,000	1,544	0.658	15.2	153	4	0	CVT	R	2	0
パレット	スズキ	1,207,500	3,884	0.658	22.5	103	4	0	CVT	R	2	1
ワゴンR	スズキ	1,212,750	13,953	0.658	25	93	4	0	CVT	R	2	1
ステラ	スバル	1,133,000	1,512	0.658	27	86	4	0	CVT	R	2	1
タント	ダイハツ	1,320,000	9,289	0.658	24.8	94	4	0	CVT	R	2	1
ミラ	ダイハツ	950,000	17,603	0.658	24.5	95	4	0	5MT	R	2	1
ムーヴ	ダイハツ	1,220,000	10,147	0.658	27	86	4	0	CVT	R	2	1
ピクシス	トヨタ	1,120,000	1,746	0.658	25.5	91	4	0	CVT	R	2	1
NBOX	ホンダ	1,240,000	2,860	0.658	22.2	104.6	4	0	CVT	R	2	1
ライフ	ホンダ	1,160,000	4,290	0.658	19.6	118.5	4	0	4AT	R	2	1
eK	三菱	982,000	2,127	0.657	21.2	110	4	0	5MT	R	2	1
モコ	日産	1,215,900	4,966	0.658	24.2	96	4	0	CVT	R	2	1
ルークス	日産	1,464,750	3,387	0.658	20.2	115	4	0	CVT	R	2	1

注：普通／小型は普通車を1に小型車を0とするダミー変数。燃料のRはレギュラー，Lは軽油，Pはプレミアム。駆動の2は2WD，4は4WD。エコ減税はエコカー減税ありが1，ない場合は0のダミー変数。通称名，メーカー，ミッション，燃料，駆動方式の5変数は名義尺度である。それ以外の8変数は連続尺度であるが，普通／小型とエコカー減税の有無は，1/0のダミー変数である。

3. DEA法による分析 [3-5, 9, 10]

重回帰分析は出力項目の目的変数が1個に限られている。これに対してDEA法は多入力多出力のデータが分析できる。しかし、入出力のデータに非負の制約があり、特に評価対象の入力データに0の値があると重みが無限大になるため、LINGOの汎用モデルでは重みに上限制約を課して対応している。また重回帰分析の変数選択のような手法がないので、用いる変数の選択に十分な注意が必要である。

しかし、CCRモデルと逆CCRモデルは、入力と出力を入れ替えたモデルと出力結果にはほぼ双対な関係がある。この特徴を用いれば、入力項目に0があり、出力項目に0がなければ、入力と出力を交換して分析することも考えられる。これを4入力（性能）2出力（価格、台数）モデルと、入力と出力を変えた2入力4出力モデルで対比し示す。

次に車の性能と価格を考えた場合、DEAは複数の機能を持つ製品の価格性能比の検討に利用できる。すなわち価格を入力として4個の機能を総合化した指標を出力に用いた1入力4出力モデルで価格性能比（総合性能／価格）を検討できる。入出力を逆にすれば、重回帰分析の結果と比較できる。前者のモデルは、価格の割に機能の最大化を試みているので、ユーザーの立場に立った製品選択の情報を提供してくれる。後者のモデルは、機能の割に高価な車種が効率的になるので、自動車メーカーの立場から製品の利益が最大化できる。

すなわち、CCRモデルで効率的なDMUに逆CCRモデルで序列を与えて改善目標を決め、「1入力固定改善法」で改善方法を具体的に提案し、東京都の公立図書館では成果を得た。しかし、製品の価格性能比は、個別の製品の改善量を求めたとしても製品全体のコンセプトにかかわるので、この方法論は部外者が利用するのは疑問視されるので行わない。価格の代わりに販売台数を用いると、前者は台数が同水準で総合機能が優れている車種が探せる。後者のモデルは、機能の割に販売戦略がうまくいって販売台数が伸びている車種を浮き彫りにできる。すなわちCCRモデルと逆CCRモデルは、異なった立場の評価を対比できる。

本稿では、DEA法の利点を総合化された入力と出力の比の最大化／最小化問題の解決に適していると考え、個別に比尺度を検討する問題点を改善する手法と考える。そして、統計手法で比較検証する。

3.1. 4機能と価格と販売台数の分析

(1) 4機能と価格と台数を出力とした分析

表3.1は4機能を入力として価格と台数を2出力とするCCRと逆CCRモデルの分析結果である。データは各変数の最大値を1以上10未満になるような単位を選んで変換してある。「SCORE」はCCRモデルのDEA効率値であり、「逆」は逆CCRモデルのDEA効率値である。重みやクロス効率値の検討は、当事者でないで行わない。CCRモデルは、総合化された性

能を入力とし、価格と台数の最大化を図っている。同じ性能水準であっても出力が最大化できる車種が効率的であるので、企業側にとって望ましいことになる。効率的なのはトヨタのアルファード、クラウン、プリウスの3車種と、スズキのワゴンR、ダイハツのタントとミラ、日産のルークスの7車種が効率的である。これらが営業的な主力車種かどうかが問題になる。そしてCCRでDEA効率値が0.87のカローラの逆CCRの効率値が1.89と最大になる。効率的な7車種の中ではクラウンの1.44が大きい。一般的にCCRモデルで効率的な車種に、逆CCRモデルのDEA非効率値の最大なものを選ぶことにしているが、この分析では効率的でないカローラが選ばれた。

表3.1 4入力2出力の分析結果

SN	通称名	メーカー	定員	CO2	燃費	排気量	価格	台数	Score	逆
1	スイフト	スズキ	5	0.93	2.50	1.24	1.24	0.2628	0.53	1.12
2	ソリオ	スズキ	5	0.95	2.45	1.24	1.48	0.2778	0.62	1.34
3	インプレッサ	スバル	5	1.32	1.76	1.6	1.71	0.2496	0.58	1.21
4	レガシイ	スバル	5	1.66	1.40	2.46	2.36	0.1742	0.57	1.00
5	アルファード	トヨタ	7	1.37	1.70	2.36	4.43	0.2535	1.00	1.42
6	ウィッシュ	トヨタ	7	1.42	1.64	1.8	2.09	0.2493	0.62	1.11
7	ヴィッツ	トヨタ	5	1.06	2.18	1.33	1.35	0.775	0.62	1.17
8	ヴォクシー	トヨタ	7	1.71	1.36	1.99	2.33	0.4678	0.76	1.36
9	ヴォルファイア	トヨタ	7	1.37	1.70	2.36	3.95	0.3762	0.94	1.65
10	エスティマ	トヨタ	7	1.87	1.24	2.36	2.96	0.387	0.83	1.39
11	カムリ	トヨタ	5	0.99	2.34	2.49	3.04	0.1499	0.94	1.00
12	カローラ	トヨタ	5	1.16	2.00	1.5	2.44	0.4771	<u>0.87</u>	<u>1.89</u>
13	クラウン	トヨタ	5	1.66	1.40	3.46	5.40	0.319	<u>1.00</u>	<u>1.44</u>
14	シエンタ	トヨタ	7	1.35	1.72	1.5	1.66	0.2418	0.58	1.03
15	ノア	トヨタ	7	1.71	1.36	1.99	2.52	0.2877	0.74	1.24
16	パッソ	トヨタ	5	1.12	2.08	1	1.00	0.4098	0.52	1.00
17	プリウス	トヨタ	5	0.71	3.26	1.8	2.17	2.2998	1.00	1.40
18	ラクティス	トヨタ	5	1.26	1.84	1.33	1.44	0.3277	0.57	1.18
19	エクストレイル	日産	5	1.85	1.42	2	2.61	0.1828	0.74	1.05
20	エルグランド	日産	7	2.15	1.08	2.49	3.61	0.1517	0.91	1.00
21	キューブ	日産	5	1.29	1.80	1.5	1.53	0.2969	0.55	1.18
22	ジューク	日産	5	1.35	1.72	1.5	1.62	0.2642	0.57	1.20
23	セレナ	日産	8	1.59	1.46	2	2.50	0.6512	0.87	1.45
24	ティーダ	日産	5	1.16	1.80	1.5	1.60	0.2169	0.56	1.19
25	ノート	日産	5	1.29	1.80	1.5	1.34	0.3656	0.50	1.03
26	マーチ	日産	5	1.03	2.26	1.2	1.04	0.1727	0.44	1.00
27	ステップワゴン	ホンダ	8	1.73	1.34	2	2.09	0.4396	0.70	1.20
28	フィット	ホンダ	5	0.89	2.60	1.34	1.59	1.5248	0.86	1.35
29	フリード	ホンダ	6	1.08	2.16	1.5	2.30	0.7817	0.85	1.77
30	MRワゴン	スズキ	4	0.86	2.70	0.66	1.26	0.1868	0.91	1.12
31	アルト	スズキ	4	0.73	3.20	0.66	1.03	0.693	0.83	1.00
32	エブリイワゴン	スズキ	4	1.53	1.52	0.66	1.34	0.1544	0.97	1.00
33	パレット	スズキ	4	1.03	2.25	0.66	1.21	0.3884	0.86	1.24
34	ワゴンR	スズキ	4	0.93	2.50	0.66	1.21	1.3953	1.00	1.26

35	ステラ	スバル	4	0.86	2.70	0.66	1.13	0.1512	0.82	1.00
36	タント	ダイハツ	4	0.94	2.48	0.66	1.32	0.9289	1.00	1.35
37	ミラ	ダイハツ	4	0.95	2.45	0.66	0.95	1.7603	1.00	1.00
38	ムーヴ	ダイハツ	4	0.86	2.70	0.66	1.22	1.0147	0.97	1.24
39	ピクシス	トヨタ	4	0.91	2.55	0.66	1.12	0.1746	0.80	1.03
40	NBOX	ホンダ	4	1.05	2.22	0.66	1.24	0.286	0.87	1.25
41	ライフ	ホンダ	4	1.19	1.96	0.66	1.16	0.429	0.83	1.10
42	eK	三菱	4	1.10	2.12	0.66	0.98	0.2127	0.68	1.00
43	モコ	日産	4	0.96	2.42	0.66	1.22	0.4966	0.89	1.24
44	ルークス	日産	4	1.15	2.02	0.66	1.46	0.3387	1.00	1.43

(2) 価格と台数を入力とし4機能を出力とした分析

表3.2は価格と台数を入力として4機能を出力とするCCRと逆CCRモデルの分析結果である。CCRモデルは、価格と台数を入力とし、総合化された性能の最大化を図っている。企業にとって同じ売り上げ水準の車種であっても、性能を最大化する車種が効率的であるとしているので、ユーザー側の視点に立って望ましい車種になる。効率的なのはスバルのレガシ、トヨタのカムリ、パッソ、日産のエルグランド、マーチと、スズキのアルト、エブリワゴン、ステラ、ダイハツのミラ、三菱のeKの10車種が効率的である。この中で5車種が小型車であり、販売額(=価格*台数)が少なく機能が良いことで選ばれたと考えられ、上で述べた解釈の正しさを裏付けている。そして逆CCRでDEA効率値が1のマーチがDEA非効率値が2.25と最大になる。また、表3.1と表3.2のCCRと逆CCRで1になるものが入れ替わっている。これはMAX=B/A;はMIN=A/B;と交換可能なためである。このため入力に0があり出力になければ、入出力を入れ替えて分析することも考えられる。

表3.2 価格と台数を入力として4機能を出力とする分析結果

SN	通称名	台数	価格	定員	CO2	燃費	排気量	Score	逆
1	スイフト	0.26	1.24	5	0.93	2.50	1.24	0.89	1.89
2	ソリオ	0.28	1.48	5	0.95	2.45	1.24	0.75	1.60
3	インプレッサ	0.25	1.71	5	1.32	1.76	1.6	0.83	1.73
4	レガシィ	0.17	2.36	5	1.66	1.40	2.46	1.00	1.74
5	アルファード	0.25	4.43	7	1.37	1.70	2.36	0.70	1.00
6	ウィッシュ	0.25	2.09	7	1.42	1.64	1.8	0.90	1.61
7	ヴェッツ	0.78	1.35	5	1.06	2.18	1.33	0.85	1.63
8	ヴォクシー	0.47	2.33	7	1.71	1.36	1.99	0.74	1.32
9	ヴォルファイア	0.38	3.95	7	1.37	1.70	2.36	0.60	1.07
10	エステイマ	0.39	2.96	7	1.87	1.24	2.36	0.72	1.20
11	カムリ	0.15	3.04	5	0.99	2.34	2.49	1.00	1.06
12	カロラ	0.48	2.44	5	1.16	2.00	1.5	0.53	1.14
13	クラウン	0.32	5.40	5	1.66	1.40	3.46	0.70	1.00
14	シエンタ	0.24	1.66	7	1.35	1.72	1.5	0.97	1.72
15	ノア	0.29	2.52	7	1.71	1.36	1.99	0.80	1.35

16	パッツ	0.41	1.00	5	1.12	2.08	1	1.00	1.92
17	プリウス	2.3	2.17	5	0.71	3.26	1.8	0.72	1.00
18	ラクティス	0.33	1.44	5	1.26	1.84	1.33	0.85	1.76
19	エクストレイル	0.18	2.61	5	1.85	1.42	2	0.95	1.36
20	エルグランド	0.15	3.61	7	2.15	1.08	2.49	1.00	1.10
21	キューブ	0.3	1.53	5	1.29	1.80	1.5	0.85	1.83
22	ジューク	0.26	1.62	5	1.35	1.72	1.5	0.83	1.74
23	セレナ	0.65	2.50	8	1.59	1.46	2	0.69	1.15
24	ティーダ	0.22	1.60	5	1.16	1.80	1.5	0.84	1.78
25	ノート	0.37	1.34	5	1.29	1.80	1.5	0.97	2.02
26	マーチ	0.17	1.04	5	1.03	2.26	1.2	<u>1.00</u>	<u>2.25</u>
27	ステップワゴン	0.44	2.09	8	1.73	1.34	2	0.83	1.43
28	フィット	1.52	1.59	5	0.89	2.60	1.34	0.74	1.16
29	フリード	0.78	2.30	6	1.08	2.16	1.5	0.56	1.17
30	MRワゴン	0.19	1.26	4	0.86	2.70	0.66	0.90	1.10
31	アルト	0.69	1.03	4	0.73	3.20	0.66	1.00	1.20
32	エブリイワゴン	0.15	1.34	4	1.53	1.52	0.66	1.00	1.03
33	パレット	0.39	1.21	4	1.03	2.25	0.66	0.81	1.17
34	ワゴンR	1.4	1.21	4	0.93	2.50	0.66	0.80	1.00
35	ステラ	0.15	1.13	4	0.86	2.70	0.66	1.00	1.22
36	タント	0.93	1.32	4	0.94	2.48	0.66	0.74	1.00
37	ミラ	1.76	0.95	4	0.95	2.45	0.66	1.00	1.00
38	ムーヴ	1.01	1.22	4	0.86	2.70	0.66	0.80	1.03
39	ピクシス	0.17	1.12	4	0.91	2.55	0.66	0.97	1.24
40	NBOX	0.29	1.24	4	1.05	2.22	0.66	0.80	1.15
41	ライフ	0.43	1.16	4	1.19	1.96	0.66	0.91	1.20
42	eK	0.21	0.98	4	1.10	2.12	0.66	1.00	1.47
43	モコ	0.5	1.22	4	0.96	2.42	0.66	0.80	1.13
44	ルークス	0.34	1.46	4	1.15	2.02	0.66	0.70	1.00

3.2. 4機能と価格の分析

(1) 4機能と価格を出力とした分析

表3.3は4機能を入力として価格を1出力とするCCRと逆CCRモデルの分析結果である。CCRモデルは、総合化された性能を入力とし、価格の最大化を図っている。後で価格を目的変数とした重回帰分析と比較する。同じ性能であっても出力の価格が最大化できる車種が効率的であるので、企業側にとって望ましいことになる。効率的なのはトヨタのアルファード、クラウンと、日産のルークスの3車種が効率的である。台数を省いて価格だけにすることで、小型車3車種が非効率になった。そして逆CCRでDEA効率値が1のアルファードのDEA非効率値が2.17と最大になり、次にクラウンが1.81である。

表3.3 4入力と価格の分析

SN	通称名	メーカー	定員	CO2	燃費	排気量	価格	Score	逆
1	スイフト	スズキ	5	0.93	2.50	1.24	1.24	0.52	1.12
2	ソリオ	スズキ	5	0.95	2.45	1.24	1.48	0.62	1.34
3	インプレッサ	スバル	5	1.32	1.76	1.6	1.71	0.57	1.24
4	レガシイ	スバル	5	1.66	1.40	2.46	2.36	0.57	1.11
5	アルファード	トヨタ	7	1.37	1.70	2.36	4.43	<u>1.00</u>	<u>2.17</u>
6	ウィッシュ	トヨタ	7	1.42	1.64	1.8	2.09	0.61	1.34
7	ヴィッツ	トヨタ	5	1.06	2.18	1.33	1.35	0.52	1.17
8	ヴォクシー	トヨタ	7	1.71	1.36	1.99	2.33	0.64	1.36
9	ヴォルファイア	トヨタ	7	1.37	1.70	2.36	3.95	0.89	1.93
10	エスティマ	トヨタ	7	1.87	1.24	2.36	2.96	0.75	1.45
11	カムリ	トヨタ	5	0.99	2.34	2.49	3.04	0.94	1.41
12	カローラ	トヨタ	5	1.16	2.00	1.5	2.44	0.85	1.89
13	クラウン	トヨタ	5	1.66	1.40	3.46	5.40	<u>1.00</u>	<u>1.81</u>
14	シエンタ	トヨタ	7	1.35	1.72	1.5	1.66	0.57	1.17
15	ノア	トヨタ	7	1.71	1.36	1.99	2.52	0.69	1.47
16	パッソ	トヨタ	5	1.12	2.08	1	1.00	0.49	1.00
17	プリウス	トヨタ	5	0.71	3.26	1.8	2.17	0.94	1.40
18	ラクティス	トヨタ	5	1.26	1.84	1.33	1.44	0.55	1.19
19	エクストレイル	日産	5	1.85	1.42	2	2.61	0.73	1.45
20	エルグランド	日産	7	2.15	1.08	2.49	3.61	0.91	1.67
21	キューブ	日産	5	1.29	1.80	1.5	1.53	0.53	1.18
22	ジューク	日産	5	1.35	1.72	1.5	1.62	0.56	1.22
23	セレナ	日産	8	1.59	1.46	2	2.50	0.67	1.45
24	ティーダ	日産	5	1.16	1.80	1.5	1.60	0.56	1.23
25	ノート	日産	5	1.29	1.80	1.5	1.34	0.47	1.03
26	マーチ	日産	5	1.03	2.26	1.2	1.04	0.44	1.00
27	ステップワゴン	ホンダ	8	1.73	1.34	2	2.09	0.57	1.20
28	フィット	ホンダ	5	0.89	2.60	1.34	1.59	0.62	1.35
29	フリード	ホンダ	6	1.08	2.16	1.5	2.30	0.80	1.77
30	MRワゴン	スズキ	4	0.86	2.70	0.66	1.26	0.91	1.29
31	アルト	スズキ	4	0.73	3.20	0.66	1.03	0.77	1.00
32	エブリイワゴン	スズキ	4	1.53	1.52	0.66	1.34	0.97	1.00
33	パレット	スズキ	4	1.03	2.25	0.66	1.21	0.85	1.25
34	ワゴンR	スズキ	4	0.93	2.50	0.66	1.21	0.87	1.27
35	ステラ	スバル	4	0.86	2.70	0.66	1.13	0.82	1.17
36	タント	ダイハツ	4	0.94	2.48	0.66	1.32	0.94	1.39
37	ミラ	ダイハツ	4	0.95	2.45	0.66	0.95	0.68	1.00
38	ムーヴ	ダイハツ	4	0.86	2.70	0.66	1.22	0.89	1.25
39	ビクシス	トヨタ	4	0.91	2.55	0.66	1.12	0.80	1.17
40	NBOX	ホンダ	4	1.05	2.22	0.66	1.24	0.87	1.28
41	ライフ	ホンダ	4	1.19	1.96	0.66	1.16	0.80	1.10
42	eK	三菱	4	1.10	2.12	0.66	0.98	0.68	1.00
43	モコ	日産	4	0.96	2.42	0.66	1.22	0.86	1.28
44	ルークス	日産	4	1.15	2.02	0.66	1.46	1.00	1.43

(2) 価格を入力とし4機能を出力とした分析

表3.4は価格を入力として4機能を出力とするCCRと逆CCRモデルの分析結果である。

CCRモデルは、価格を入力とし、総合化された性能の最大化を図っているため、製品の価格性能比の評価に今後利用できると思われる。同じ価格帯の車種であっても、性能を最大化する車種が効率的であるため、ユーザー側の視点に立って望ましい車種になる。効率的なのはトヨタのパッソ、日産のマーチと、スズキのアルトが効率的である。表3.2に比べて普通車の3車種、小型車の4車種が非効率になった。そして逆CCRでDEA効率値が1のマーチのDEA非効率値が2.32と最大になる。

表3.4 価格を入力として4機能を出力とする分析結果

SN	通称名	価格	定員	CO2	燃費	排気量	Score	逆
1	スイフト	1.24	5	0.93	2.50	1.24	0.89	1.92
2	ソリオ	1.48	5	0.95	2.45	1.24	0.74	1.62
3	インプレッサ	1.71	5	1.32	1.76	1.6	0.81	1.77
4	レガシイ	2.36	5	1.66	1.40	2.46	0.90	1.75
5	アルファード	4.43	7	1.37	1.70	2.36	0.46	1.00
6	ウィッシュ	2.09	7	1.42	1.64	1.8	0.74	1.64
7	ヴィッツ	1.35	5	1.06	2.18	1.33	0.85	1.91
8	ヴォクシー	2.33	7	1.71	1.36	1.99	0.74	1.57
9	ヴォルファイア	3.95	7	1.37	1.70	2.36	0.52	1.12
10	エステイマ	2.96	7	1.87	1.24	2.36	0.69	1.34
11	カムリ	3.04	5	0.99	2.34	2.49	0.71	1.06
12	カローラ	2.44	5	1.16	2.00	1.5	0.53	1.18
13	クラウン	5.40	5	1.66	1.40	3.46	0.55	1.00
14	シエンタ	1.66	7	1.35	1.72	1.5	0.86	1.75
15	ノア	2.52	7	1.71	1.36	1.99	0.68	1.45
16	パッソ	1.00	5	1.12	2.08	1	1.00	<u>2.03</u>
17	プリウス	2.17	5	0.71	3.26	1.8	0.72	1.06
18	ラクティス	1.44	5	1.26	1.84	1.33	0.80	1.81
19	エクストレイル	2.61	5	1.85	1.42	2	0.66	1.36
20	エルグランド	3.61	7	2.15	1.08	2.49	0.60	1.10
21	キューブ	1.53	5	1.29	1.80	1.5	0.85	1.88
22	ジューク	1.62	5	1.35	1.72	1.5	0.80	1.77
23	セレナ	2.50	8	1.59	1.46	2	0.69	1.50
24	ティーダ	1.60	5	1.16	1.80	1.5	0.81	1.80
25	ノート	1.34	5	1.29	1.80	1.5	0.97	2.14
26	マーチ	1.04	5	1.03	2.26	1.2	1.00	<u>2.32</u>
27	ステップワゴン	2.09	8	1.73	1.34	2	0.83	1.75
28	フィット	1.59	5	0.89	2.60	1.34	0.74	1.60
29	フリード	2.30	6	1.08	2.16	1.5	0.56	1.25
30	MRワゴン	1.26	4	0.86	2.70	0.66	0.76	1.10
31	アルト	1.03	4	0.73	3.20	0.66	1.00	<u>1.30</u>
32	エブリイワゴン	1.34	4	1.53	1.52	0.66	0.87	1.03
33	パレット	1.21	4	1.03	2.25	0.66	0.76	1.18
34	ワゴンR	1.21	4	0.93	2.50	0.66	0.77	1.15
35	ステラ	1.13	4	0.86	2.70	0.66	0.85	1.22
36	タント	1.32	4	0.94	2.48	0.66	0.71	1.06
37	ミラ	0.95	4	0.95	2.45	0.66	0.98	<u>1.48</u>

38	ムーヴ	1.22	4	0.86	2.70	0.66	0.79	1.13
39	ピクシス	1.12	4	0.91	2.55	0.66	0.84	1.24
40	NBOX	1.24	4	1.05	2.22	0.66	0.74	1.16
41	ライフ	1.16	4	1.19	1.96	0.66	0.78	1.25
42	eK	0.98	4	1.10	2.12	0.66	0.93	1.47
43	モコ	1.22	4	0.96	2.42	0.66	0.76	1.16
44	ルークス	1.46	4	1.15	2.02	0.66	0.62	1.00

3.3. 4機能と台数の分析

(1) 4機能と台数を出力とした分析

表3.5は4機能を入力として台数を出力とするCCRと逆CCRモデルの分析結果である。CCRモデルは、総合化された性能を入力とし、台数の最大化を図っている。後で台数を目的変数とした重回帰分析と比較する。同じ性能であっても台数が最大化できる車種が効率的であるので、企業側にとって望ましいことになる。効率的なのはトヨタのプリウスと、ダイハツのミラである。そして逆CCRでミラが12.08と最大になる。

表3.5 4入力と台数の出力モデル

SN	通称名	メーカー	定員	CO2	燃費	排気量	台数	Score	逆
1	スイフト	スズキ	5	0.93	2.50	1.242	0.2628	0.15	1.69
2	ソリオ	スズキ	5	0.95	2.45	1.242	0.2778	0.16	1.80
3	インプレッサ	スバル	5	1.32	1.76	1.599	0.2496	0.20	1.81
4	レガシィ	スバル	5	1.66	1.40	2.457	0.1742	0.17	1.17
5	アルファード	トヨタ	7	1.37	1.70	2.362	0.2535	0.21	1.51
6	ウィッシュ	トヨタ	7	1.42	1.64	1.797	0.2493	0.21	1.50
7	ヴィッツ	トヨタ	5	1.06	2.18	1.329	0.775	0.49	5.31
8	ヴォクシー	トヨタ	7	1.71	1.36	1.986	0.4678	0.48	2.94
9	ヴォルファイア	トヨタ	7	1.37	1.70	2.362	0.3762	0.31	2.24
10	エステイマ	トヨタ	7	1.87	1.24	2.362	0.387	0.43	2.48
11	カムリ	トヨタ	5	0.99	2.34	2.493	0.1499	0.09	1.00
12	カローラ	トヨタ	5	1.16	2.00	1.496	0.4771	0.33	3.35
13	クラウン	トヨタ	5	1.66	1.40	3.456	0.319	0.32	1.53
14	シエンタ	トヨタ	7	1.35	1.72	1.496	0.2418	0.20	1.43
15	ノア	トヨタ	7	1.71	1.36	1.986	0.2877	0.29	1.81
16	パッソ	トヨタ	5	1.12	2.08	0.996	0.4098	0.27	2.84
17	プリウス	トヨタ	5	0.71	3.26	1.797	2.2998	1.00	11.95
18	ラクティス	トヨタ	5	1.26	1.84	1.329	0.3277	0.25	2.35
19	エクストレイル	日産	5	1.85	1.42	1.995	0.1828	0.18	1.22
20	エルグランド	日産	7	2.15	1.08	2.488	0.1517	0.20	1.00
21	キューブ	日産	5	1.29	1.80	1.498	0.2969	0.23	2.14
22	ジューク	日産	5	1.35	1.72	1.498	0.2642	0.21	1.92
23	セレナ	日産	8	1.59	1.46	1.997	0.6512	0.62	3.63
24	ティーダ	日産	5	1.16	1.80	1.498	0.2169	0.17	1.60
25	ノート	日産	5	1.29	1.80	1.498	0.3656	0.28	2.64
26	マーチ	日産	5	1.03	2.26	1.198	0.1727	0.11	1.17

27	ステップワゴン	ホンダ	8	1.73	1.34	1.997	0.4396	0.46	2.50
28	フィット	ホンダ	5	0.89	2.60	1.339	1.5248	0.82	9.55
29	フリード	ホンダ	6	1.08	2.16	1.496	0.7817	0.50	4.80
30	MRワゴン	スズキ	4	0.86	2.70	0.658	0.1868	0.11	1.24
31	アルト	スズキ	4	0.73	3.20	0.658	0.693	0.47	3.87
32	エブリイワゴン	スズキ	4	1.53	1.52	0.658	0.1544	0.14	1.12
33	パレット	スズキ	4	1.03	2.25	0.658	0.3884	0.24	2.74
34	ワゴンR	スズキ	4	0.93	2.50	0.658	1.3953	0.80	9.51
35	ステラ	スバル	4	0.86	2.70	0.658	0.1512	0.09	1.00
36	タント	ダイハツ	4	0.94	2.48	0.658	0.9289	0.53	6.34
37	ミラ	ダイハツ	4	0.95	2.45	0.658	1.7603	1.00	<u>12.08</u>
38	ムーヴ	ダイハツ	4	0.86	2.70	0.658	1.0147	0.62	6.71
39	ピクシス	トヨタ	4	0.91	2.55	0.658	0.1746	0.10	1.18
40	NBOX	ホンダ	4	1.05	2.22	0.658	0.286	0.18	2.02
41	ライフ	ホンダ	4	1.19	1.96	0.658	0.429	0.30	3.11
42	eK	三菱	4	1.10	2.12	0.657	0.2127	0.14	1.52
43	モコ	日産	4	0.96	2.42	0.658	0.4966	0.29	3.42
44	ルークス	日産	4	1.15	2.02	0.658	0.3387	0.23	2.44

(2) 台数を入力とし4機能を出力とした分析

表3.6は台数を入力として4機能を出力とするCCRと逆CCRモデルの分析結果である。CCRモデルは、台数を入力とし、総合化された性能の最大化を図っているため、販売台数は少ないが高性能な車種が効率的であるため、ユーザー側の視点に立って望ましい車種になる。効率的なのはトヨタのカムリ、日産のエルグランデと、スバルのステラである。そして逆CCRのDEA非効率値はカムリが11.22と最大になる。

表3.6 台数を入力として4機能を出力とする分析結果

SN	通称名	台数	定員	CO2	燃費	排気量	Score	逆
1	スイフト	0.26	5	0.93	2.50	1.24	0.59	6.82
2	ソリオ	0.28	5	0.95	2.45	1.24	0.55	6.34
3	インプレッサ	0.25	5	1.32	1.76	1.6	0.55	5.07
4	レガシィ	0.17	5	1.66	1.40	2.46	0.86	5.77
5	アルファード	0.25	7	1.37	1.70	2.36	0.66	4.82
6	ウィッシュ	0.25	7	1.42	1.64	1.8	0.67	4.73
7	ヴィッツ	0.78	5	1.06	2.18	1.33	0.19	2.02
8	ヴォクシー	0.47	7	1.71	1.36	1.99	0.34	2.09
9	ヴォルファイア	0.38	7	1.37	1.70	2.36	0.45	3.25
10	エスティマ	0.39	7	1.87	1.24	2.36	0.40	2.30
11	カムリ	0.15	5	0.99	2.34	2.49	1.00	<u>11.22</u>
12	カローラ	0.48	5	1.16	2.00	1.5	0.30	3.01
13	クラウン	0.32	5	1.66	1.40	3.46	0.65	3.15
14	シエンタ	0.24	7	1.35	1.72	1.5	0.70	5.11
15	ノア	0.29	7	1.71	1.36	1.99	0.55	3.40
16	パッツ	0.41	5	1.12	2.08	1	0.35	3.65
17	プリウス	2.3	5	0.71	3.26	1.8	0.08	1.00

18	ラクティス	0.33	5	1.26	1.84	1.33	0.42	4.03
19	エクストレイル	0.18	5	1.85	1.42	2	0.82	5.58
20	エルグランド	0.15	7	2.15	1.08	2.49	1.00	5.12
21	キューブ	0.3	5	1.29	1.80	1.5	0.47	4.36
22	ジューク	0.26	5	1.35	1.72	1.5	0.52	4.68
23	セレナ	0.65	8	1.59	1.46	2	0.28	1.61
24	ティーダ	0.22	5	1.16	1.80	1.5	0.63	5.96
25	ノート	0.37	5	1.29	1.80	1.5	0.38	3.54
26	マーチ	0.17	5	1.03	2.26	1.2	0.86	9.40
27	ステップワゴン	0.44	8	1.73	1.34	2	0.40	2.19
28	フィット	1.52	5	0.89	2.60	1.34	0.10	1.22
29	フリード	0.78	6	1.08	2.16	1.5	0.21	1.99
30	MRワゴン	0.19	4	0.86	2.70	0.66	0.81	8.78
31	アルト	0.69	4	0.73	3.20	0.66	0.26	2.12
32	エブリイワゴン	0.15	4	1.53	1.52	0.66	0.89	7.07
33	パレット	0.39	4	1.03	2.25	0.66	0.37	4.16
34	ワゴンR	1.4	4	0.93	2.50	0.66	0.11	1.24
35	ステラ	0.15	4	0.86	2.70	0.66	1.00	10.85
36	タント	0.93	4	0.94	2.48	0.66	0.16	1.88
37	ミラ	1.76	4	0.95	2.45	0.66	0.08	1.00
38	ムーヴ	1.01	4	0.86	2.70	0.66	0.15	1.62
39	ピクシス	0.17	4	0.91	2.55	0.66	0.85	9.78
40	NBOX	0.29	4	1.05	2.22	0.66	0.49	5.58
41	ライフ	0.43	4	1.19	1.96	0.66	0.32	3.28
42	eK	0.21	4	1.10	2.12	0.66	0.66	7.16
43	モコ	0.5	4	0.96	2.42	0.66	0.29	3.50
44	ルークス	0.34	4	1.15	2.02	0.66	0.41	4.29

3.4. DEAから分かること

DEAから分かった効率的な車種を表3.7にまとめる。価格と台数を出力とすると、企業側の評価になる。7車種が選ばれる。カローラは効率的でないが、逆CCRでDEA非効率値が最大になる。価格だけを出力にすると7車種のうち3車種が選ばれ、アルファードの非効率値が最大になる。台数を出力にすると、2車種が選ばれ、ミラがDEA非効率値が最大になる。価格と台数別の分析では効率的な車種が異なる。

価格と台数を入力にすると、ユーザー側の視点にたった評価と考えられる。表3.2の10車種が選ばれ、普通車と小型車が5車種ずつになる。価格だけを入力にすると10車種のうち3車種が選ばれる。いずれもマーチのDEA非効率値が最大になる。台数だけを入力にすると別の3車種が選ばれ、カムのDEA非効率値が最大になる。

ミラは表3.1と表3.5に含まれるのは、価格より台数が企業に貢献しているからと考えられる。表3.2にも含まれるのは、価格が安く台数も多いということで、ユーザー側にも企業側にも良いという特徴を持っているからと考えられる。

統計の評価では、下線を引いたDEA非効率値最大の車種に注目し比較検討する。

表3.7 DEAの効率的な車種

企業	表3.1 (価格, 台数)	アルファード, (カローラ), <u>クラウン</u> , プリウス, ワゴンR, タント, ミラ, ルークス
	表3.3 (価格)	<u>アルファード</u> , クラウン, ルークス
	表3.5 (台数)	プリウス, <u>ミラ</u>
ユーザー	表3.2 (価格, 台数)	レガシ, カムリ, パッソ, エルグランド, <u>マーチ</u> , アルト, エブリワゴン, ステラ, ミラ, eK
	表3.4 (価格)	パッソ, <u>マーチ</u> , アルト
	表3.6 (台数)	<u>カムリ</u> , エルグランド, ステラ

4. 統計分析

4.1. 一元配置の分散分析

価格と販売台数を一元配置の分散分析で分析し評価する。ただし、DEAで性能を入力として企業側に立って効率的と考えられる車種を□, その中で逆CCRで最大のDEA非効率値をもつ車種を■で表し車種名も表示する。価格と台数を入力としてユーザー側に立って効率的な車種を○, その中で逆CCRでDEA非効率値が最大値になる車種を●で表し車種名を表示する。

(1) 価格の分析

価格を5個の名義尺度で一元配置の分散分析する。図4.1は、メーカー別に分析したものである。HSD検定では、7社の車種数が少ないので平均に差がなかった。しかし、箱髭図をみるとトヨタは15車種と一番多く価格帯が他社を圧倒し、クラウンが最上位の高級車であることが分かる。ホンダと日産は重なりが大きいですが、日産の価格帯がより広い。スズキとダイハツの価格は、全車種平均の200万円以下にあることが分かる。クラウン、アルフォード、ミラが企業側の立場、マーチとカムリがユーザー側の立場で評価が高いと考えることができる。一元配置の分散分析のグラフ表現である層別箱髭図は、簡単に有効な情報が分かる。また中学数学にも箱髭図が取り入れられているので、教育ギャップをなくすために大学教育で学生に教える必要がある。また、箱髭図の元になる分位点(パーセント点)も統計の入門として重要である。

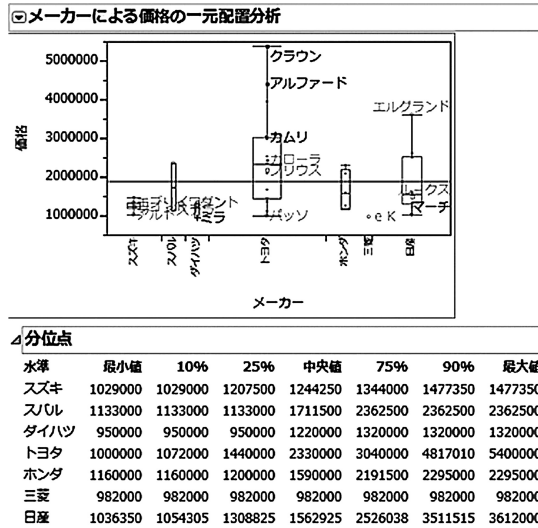


図4.1 メーカー別に価格を分析

図4.2は、駆動方式別に分析したものである。HSD検定では、2WDと4WDの平均に差があった。2WDでは、クラウンとエルグランドが外れ値になっている。4WDの価格が高いため、これらは4と2という順序尺度で分析することも考えられる。企業評価のクラウンとアルファードが表示されるがミラは2WDの下の鬚に表示されている。ユーザー評価のカムリは表示されているがマーチはミラと重なり少し上に表示されている。ただしこの項目は、各車種のカatalog性能をみれば多くの車種で両方提供されているので、今回データ収集に用いた区分であり、分析項目としては適していないと考えるべきであろう。

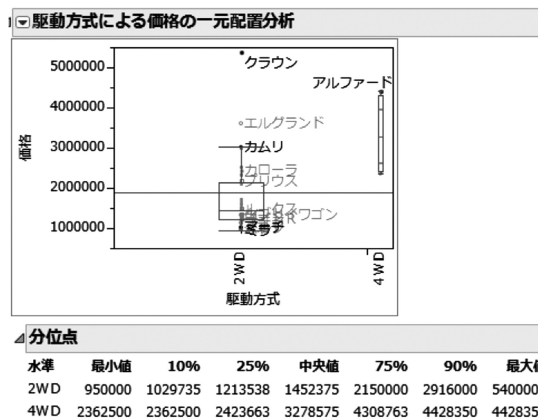


図4.2 駆動方式別に価格を分析

燃料はレギュラーが圧倒的に多いので、今回の分析に利用できない内容である。

(2) 台数の一元配置

図4.3は、メーカー別に台数を分析したものである。HSD検定では、ダイハツの販売台数が一番大きく、日産が一番少ない。残りの会社はダイハツと比べると平均に差がなく、日産と比べると差がないグループになった。

企業側の立場に立つミラはダイハツで売れている車種であるが、トヨタの企業側に立つクラウン、アルフォードはユーザー側に立つカムリと台数では中央値以下である。日産のマーチは、価格ではユーザー側の立場で評価が高いが、台数では不振であることが分かる。

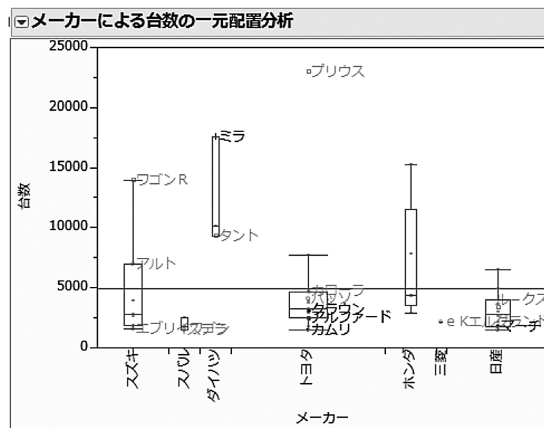


図4.3 台数のメーカーによる一元配置

駆動方式は省略する。

4.2. 価格で単回帰分析

性能の4変数を価格で単回帰分析する。図4.4の左は総排気量を目的変数とした分析である。回帰直線の下側に企業側の評価で選ばれた3車種、上側にユーザー側の評価で選ばれた2車種が布置する。すなわち価格に対して総排気量の誤差が負のものが企業側の評価で選ばれ、誤差が正のものがユーザー側の評価で選ばれる。

右は燃費を目的変数とした分析である。5車種とも回帰直線の上側に布置されている。直線の下にくる車種より、燃費効率が悪いことになる。ユーザー側に立つマーチとカムリが下側に布置されないのが不思議である。

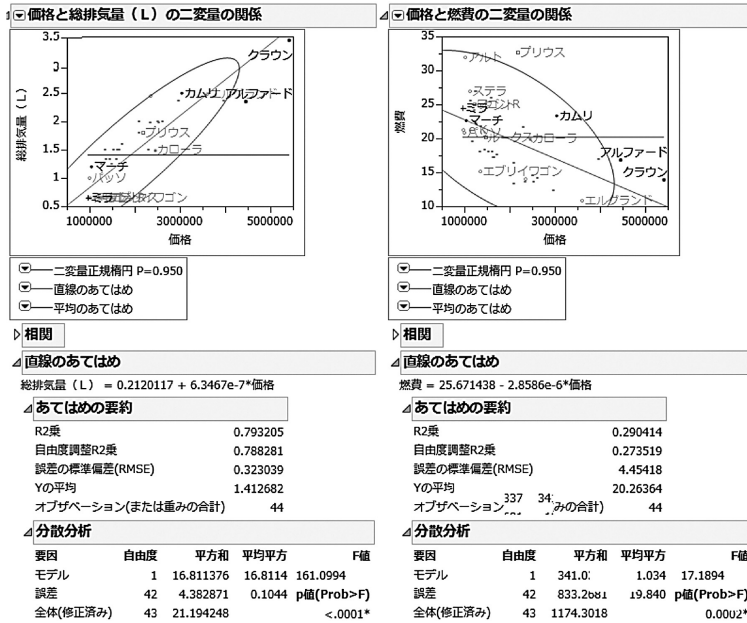


図4.4 総排気量 (左) と燃費 (右) を目的変数とした価格による単回帰分析

図4.5の左はCO2を目的変数とした分析である。回帰直線の下に5車種が布置される。企業側の評価に立つクラウン、アルフォード、ミラが回帰直線の上に布置されないのは、環境対応を重視して成功しているのか検討する必要がある。右は定員を目的変数とした分析で、マーチが回帰直線の上に来て、普通車であるため高価格なカムリは5人であるが直線の下に来ている。ただし小型車の全車種が4人であり、分析に用いることの是非は考えなければいけない。

以上のように、価格性能比を単回帰分析で個別に検討しても明確な判断ができない。DEAは、価格と総合性能比で全体的な判断が可能になる。

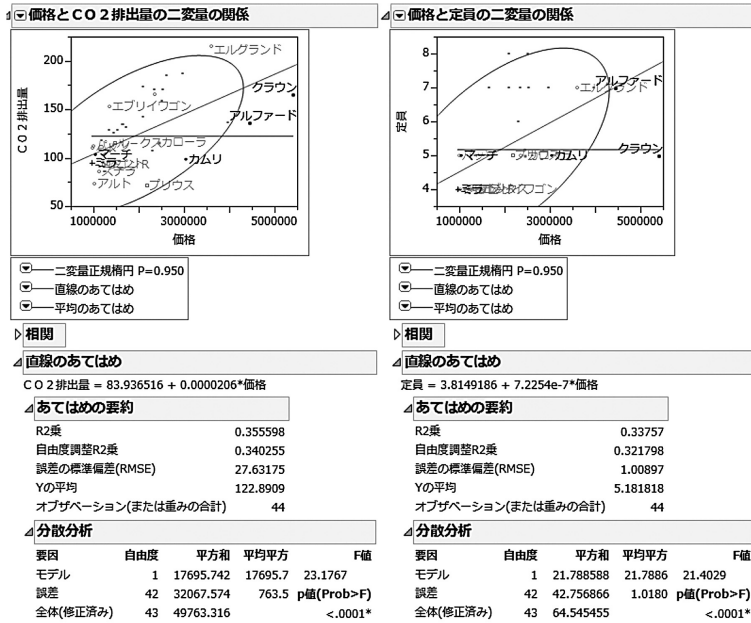


図4.5 CO2 (左) と定員 (右) を目的変数とした価格による単回帰分析

4.3. 重回帰分析

(1) 価格を性能で重回帰する

図4.6は価格を4個の性能で重回帰分析した結果である。総排気量だけが5%で棄却されるので変数選択を行う必要がある。VIF (Variance Inflation Factor) は多重共線性を診断する統計量である。燃費はほぼ10であり、燃費を残りの3変数で重回帰分析すると決定係数が0.9になることを示す。最近、間違った解説書のためか、「VIFが10ぐらいで多重共線性がある。以上で分析を終える」というレポートが散見される。基準はないが少なくとも50以上、望むらくは100以上を基準にするのが望ましい。そして、多重共線性があっても、変数選択を行えば、該当する変数がモデルから掃き出されることが多い。これは文献 [17] をみるか、筆者のCPDに関する論文が参考になる。

「予測値と残差のプロット」をみると、企業で評価される2車種は正の大きな残差、ミラは0、そしてユーザーで評価される2車種は大きな負の残差になる。DEAの特徴とされるDMUは凸体に包み込まれるという説明は、重回帰分析の「予測値と残差のプロット」でクラウン、アルト、ミラ、マーチ、カムリにほぼ包み込まれることに対応していると考えられ、DEAの際立った特徴とは考えられない。

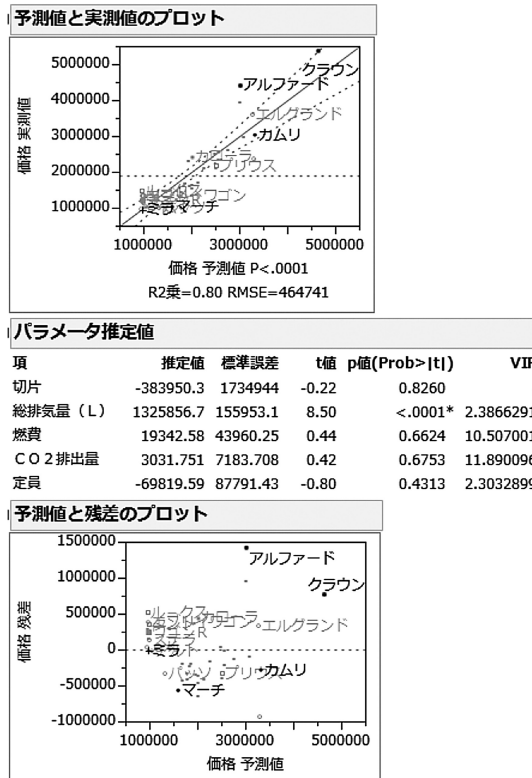


図4.6 価格を4個の性能で重回帰分析した結果

(2) 台数を性能で重回帰する

図4.7は、台数を目的変数とした重回帰分析の出力である。「残差のプロット」からミラとクラウンが正で、残りは負であるが、価格ほど明確な情報が得られない。

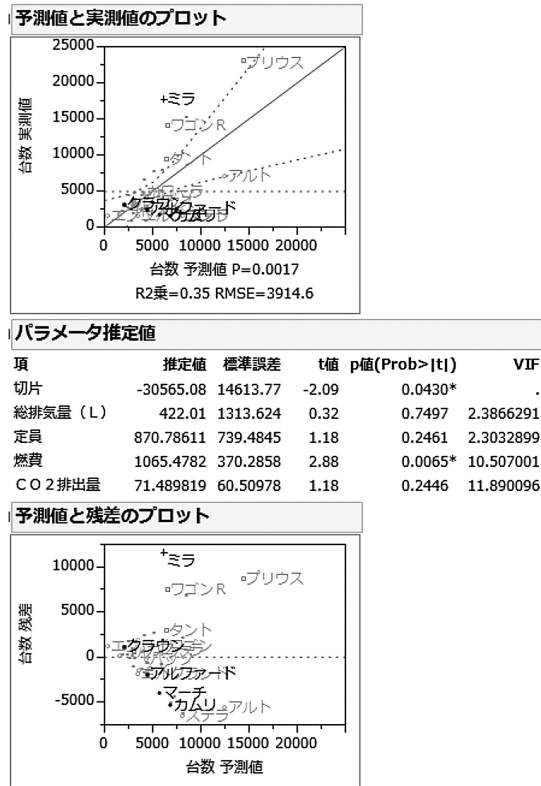


図4.7 台数の重回帰分析による予測

4.4. 主成分分析

図4.8は元の6変数を用いた主成分分析である。第2主成分までが固有値が1以上であり累積固有値が80%程度である。右の因子負荷プロットから台数と燃費は第2象限に、CO2排出量が第4象限に、それ以外の3性能は第1象限に布置される。真ん中のスコアプロットから、5車種に関する重要な傾向は読み取れない。

図4.9は元の価格と台数を入力とし残りの4個の性能との8個の1入力1出力の比を作り、主成分分析を行った。第3主成分までが固有値が1以上であり累積固有値が80%以上である。排気/価格は第1象限にくるが残りの価格を入力とする3個は第2象限にくる。0度から170度の範囲に全ての比が布置されている。この範囲がユーザー側に立って評価される車種に対応し、スコアプロットではカムリとマーチが布置される。原点と対称なところに、クラウン、アルフォード、ミラが布置され、この5車種が44車種を包み込んでいて、DEAの包絡分析の特徴を表している。すなわち、主成分分析でも入出力の比にデータを変換すれば、DEAの特徴をある程度とらえることができる。これは、東京都23区の公立図書館の25年間の違いを評

価する際に用いて成功している [13]。

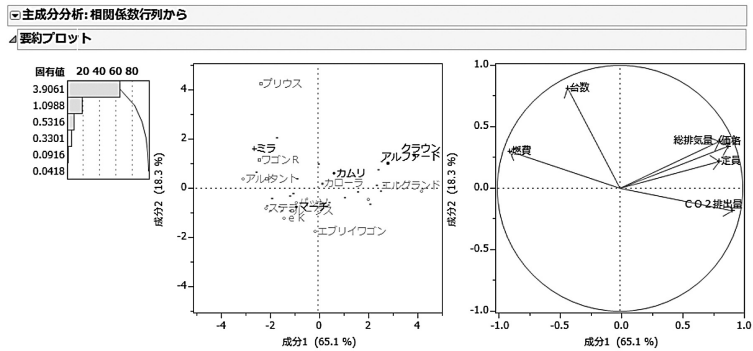


図4.8 元の6変数を用いた主成分分析

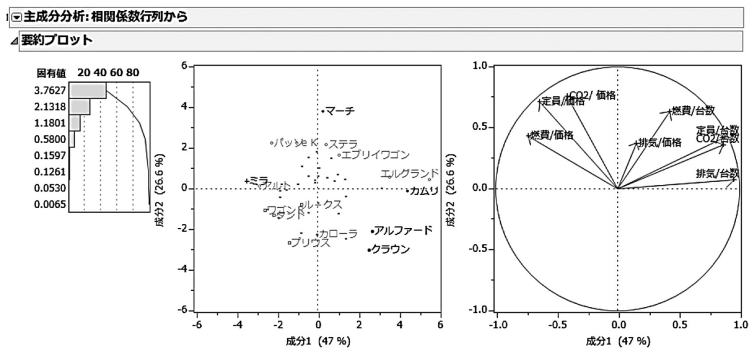


図4.9 8個の比を用いた主成分分析

4.5. クラスタ分析

図4.10はクラスタ分析のWard法で8個の比を分析した結果である。クラスタ分析で、ケースのクラスタ結果をうまく解釈できる例は少ない。今回も元のデータで分析しては分からない。しかし比をとり逆DEAモデルでDEA非効率的値が最大になる車種に注目することでケースのクラスタ分けの意味が明らかになった。上のスイフトからエステイマの16車種は普通車でありDEAでは評価されない車種である。次のパッソ、小型車のeK、マーチはユーザー側に立って評価されるパッソが含まれている。第3のMRワゴンからの4車種は小型車である。カローラからの4車種は普通車である。アルトからミラを含むルークスまでの10車種は小型車である。レガシから最後のクラウンまでの6車種にはカムリ、アルフォード、クラウンといったDEAで評価れる3車種が含まれている。しかし、主成分分析ほど明確な区

分はできない。

変数のクラスターは、排気／価格が台数との4個の比とクラスターになり、残りの3個の価格との比と別れている。

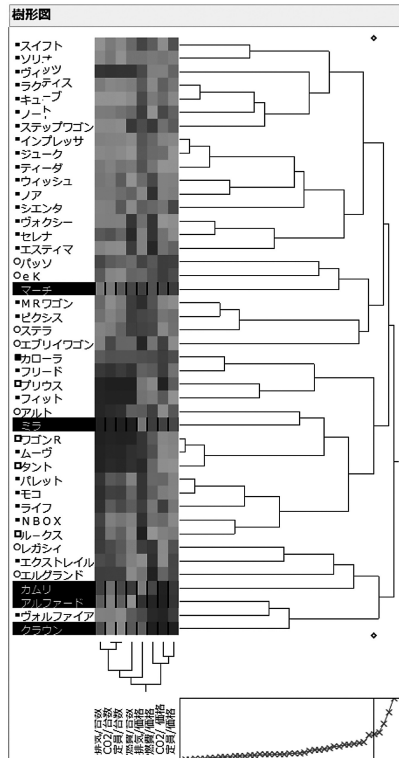


図4.10 クラスター分析

5. 普通車と小型車の判別

5.1. 分散共分散行列に基づく判別関数の終焉

1997年に始めた整数計画法による誤分類数最小化 (MNM) 基準による最適線形判別関数の研究も2010年に終了し、研究成果を著書にまとめて出版できた [8]。そこで分かったことは、①判別係数と誤分類数の関係、②MNMの単調減少性とMNM=0になる判別モデルが分かれば、その説明変数を含む全ての判別モデルがMNM=0になること、③スイス銀行紙幣データは(X4, X6)という2変数でMNM=0になり、それを含む16個の判別モデルが線形分離可能であるが既存の判別関数は線形分離可能なデータを認識できないという驚く事実、④判別関数は判別超平面上のケースがいずれの群に判別されるかという問題を未解決のまま放置

してきたが唯一改定IP-OLDFがそれを解決できた。すなわち既存の判別関数の誤分類数は正しいことが保証されていない。⑤LDFとロジスティック回帰という既存の代表的な判別関数とMNMの近似値を求める改定IPLP-OLDFを、研究に用いていた4種類の実データからリサンプリング標本を作成し、135個の異なった判別モデルで評価を行い圧倒的に良い結果を得た。

MNM基準に基づく判別関数は、教師標本で良い結果を得ても検証標本で悪くなると一般的に考えられてきた。しかし結果が出てしまえば、現実には合わない仮説（Fisherの仮説）で作上げられてきた判別理論は、それを前提にどれだけ精緻な理論をその後で構築しても検証標本で判別成績は悪い。正規分布でない多くの現実のデータで得られた判別関数を検証標本に当てはめても良いわけがないのに、ノンパラメトリックな手法である改定IP-OLDFの方がOverestimateするという間違った推論である。この間違いは、近年医学や経済で判別分析のユーザーがLDFやQDFを使わないでロジスティック回帰を利用していることから分かる。ロジスティック回帰は特定の分布を想定せず、データにあうロジスティック曲線で判別しているから、判別成績が良いだけという常識的な判断ができない研究者が多い。また、2群を正規分布と考えているがLDFやQDFでは、判別係数や誤分類確率の標準誤差が分かっているが、推測統計学と無縁な学問であることが分かっている。単に2群を正規分布とすることで、Fisherが考えた分散比最大化基準による線形判別関数が、正規分布と関係する分散共分散行列から簡単に定式化でき、統計ソフトとして実現しやすかったからである。

すなわち、Fisherは判別される2群の分散比（群間分散／群内分散）の最大化から式（1）で示される線形判別関数（LDF）を提案した。これは制約式なしの非線形計画法になるが、式（2）のように分母を定数に置き換えて制約に取り込むことで、2次計画法（Quadratic Programming, QP）で解くことができる。

$$\text{MIN}=\mathbf{b}' (x_{m1}-x_{m2}) (x_{m1}-x_{m2})' \mathbf{b}/\mathbf{b}' \sum \mathbf{b}; \quad (1)$$

$$\text{あるいは、MIN}=\mathbf{b}' (x_{m1}-x_{m2}) (x_{m1}-x_{m2})' \mathbf{b}; \quad \mathbf{b}' \sum \mathbf{b}=1; \quad (2)$$

その後、2群がFisherの仮説を満たせば、分散比最大化基準によるLDFが、容易に2群を表す正規分布関数の対数尤度比で式（3）のように定式化されることが分かった。この式の利点は、分散共分散行列 Σ から直接LDFの計算式が定式化されることで、統計ソフトに取り入れられその後の分散共分散行列による判別関数が広く普及した原因と考えられる。判別関数の評価にはLOO法がある。また判別される2群に1/-1のダミー変数を与えて形式的に重回帰分析として変数選択法などの検定統計量が利用できる。

$$\text{LDF} : f(x) = \{x - (m_1 + m_2) / 2\}' \Sigma^{-1} (m_1 - m_2) \quad (3)$$

2群の分散共分散行列が等しくない場合は、式(4)の2次判別関数(QDF)が定式化される。

$$f(x) = x' (\Sigma_2^{-1} - \Sigma_1^{-1}) x / 2 + (m_1' \Sigma_1^{-1} - m_2' \Sigma_2^{-1}) x + c \quad (4)$$

また式(5)のマハラノビスの汎距離を用いて、多群判別や品質管理のMT理論に応用されている。

$$D = \text{SQRT}((x - m)' \Sigma^{-1} (x - m)) \quad (5)$$

分散共分散行列の逆行列を計算するだけでLDFやQDFや多群判別が定式化されて理論展開しやすいが、重回帰分析と異なり推測統計学的知見が得られないため、今日では役目を終えたLOO法で学習データと検証データによる評価が必要であった。さらに分散共分散に基づく判別分析には重大な問題がある。それは分散共分散の逆行列を計算する必要があり、ある説明変数が一定値をとる場合はランク落ちして判別関数が求まらないので、次のように分類して考える必要がある。

- 1) 2群のある説明変数が同じ値である場合は、判別に役立たない変数であり、LDFもQDFも「平均値の差のt検定」もこの変数を分析から省くことは正しい。
- 2) 2群の説明変数がそれぞれ別の一定値をとる場合、この変数だけで判別できることはt検定からも分かる。しかし統計ソフトによってはLDFの誤分類数は必ずしも0にならず、QDFはランク落ちして計算できない。
- 3) 2群の一方が一定値をとり、他方がばらつく場合は、2)に次いで判別に重要な変数である。LDFは計算できるが、QDFはランク落ちする。正則化判別関数 [16] の改良アルゴリズムを採用しているJMPは、class2を全てclass1に誤判別する。

これらに対して、ある統計ソフトは1) 2) 3) に対応する変数をすべて分析から除外し、JMPは3) に対してQDFは正しく対応していないので、利用に際して注意する必要がある。この混乱は、「群に属するデータは必ずばらつき分散が0である場合を想定していない」ことが原因である。

式(6)のロジスティック回帰は、LDFやQDFより誤分類数が少ないことが経験的に知られているので、医学や経済の分野で利用されている。また、一部の例外を除いてMNM=0のデータを認識できる。

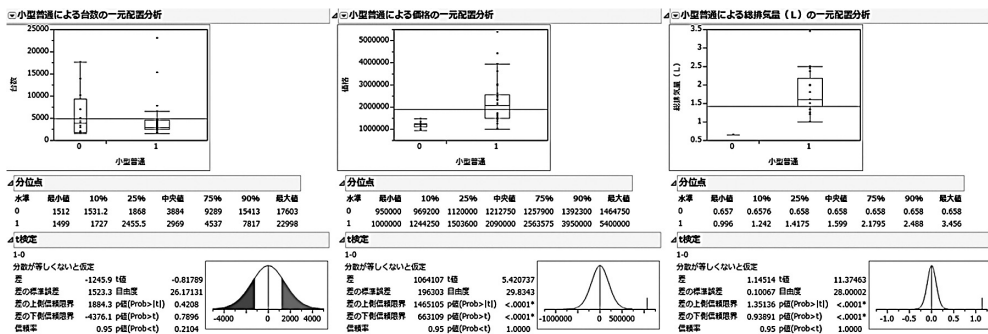
$$P=1/(1+\exp(-b_0-b_1X_1-\dots-b_kX_k)) \quad (6)$$

筆者の研究書に対して竹内啓先生が書評 [15] していただき、多くの退官された先生からおほめをいただいた。

その後、応用研究として「試験の合否判定 [11, 12, 14]」でもって、既存の判別関数が線形分離可能なデータを認識できないことの検証を2012年末まで行った。このように時間がかかったのは、大きなハプニングがあり、2010年から始めた統計入門の試験結果の研究に切り替えたためである。また、多くの試験の合否判定で合格群の全てが不合格群に誤判別されるという異常状態の解明に時間を浪費した。それが一般化逆行列と正則化判別関数という最新手法の問題であることが分かり、2012年に終了した。今回、本データを用いて説明する。

5.2. 判別分析に重要な一元配置の分散分析

合否判定を合格得点の10%、50%、90%水準の3水準で小問100問やそれをまとめた大問の得点で2群判別し、18個の合否判定でLDFとQDFの誤分類確率を求めて、科学的に判別する手法の判別関数の誤分類数が高いことが分かった。また、ある合否判定で合格群がすべて不合格群に誤判別される異常状態の解明に3年も費やした。それを解明したのが説明変数の2群による層別箱髭図である。図5.1は、左上から右下にかけて、台数、価格、総排気量、燃費、CO2排出量、定員の層別箱髭図である。台数から普通車の75%点は小型車の中央値をわずかに上回っているが、最大値は小型車より多い。小型車の価格は、普通車の25%よりも低価格である。総排気量は、小型車が0.657から0.658の範囲であり、普通車とは完全に分離している。燃費は、小型車の最小値が普通車の25%から50%の間にある。CO2排出量は、最小値がほぼ同じで小型車の75%点が普通車の25%から50%の間にある。定員は小型車が4人と一定値であり、普通車の5人以上であり、MNM=0の判別ができる。



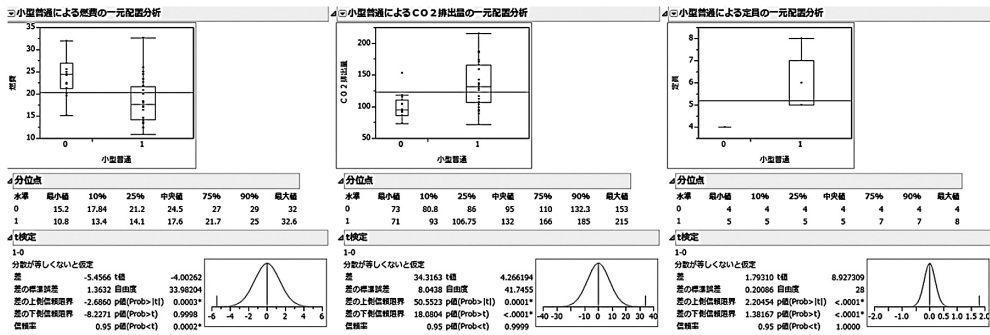


図5.1 各変数と小型車／普通車の一元配置

これに対して平均の差のt検定は、総排気量が11.37、定員が8.93、価格が5.42、CO2が4.27、燃費が-4.00、台数が-0.82であり、1変数で考えるとこの順で判別成績が悪くなる。特に、定員は小型車が一定の値をとってもt検定は影響されないが、判別関数は統計ソフトによって種々の異なった対応をとっていて利用には確認して使う必要がある。T検定から、総排出量と定員はこの1変数でMNM=0で判別できる重要な説明変数である。

5.3. 分散共分散行列に基づく判別関数の終焉

T検定から、総排出量と定員でMNM=0で判別できることが分かる。表5.1は普通車に1、小型車に0のダミー変数を与えて目的変数とし、6変数で変数増加法を行い変数欄の順に説明変数が取り込まれた。変数減少法では6変数のフルモデルから台数、燃費の順に掃き出される。Cp統計量は6変数、AICとBICから5変数のモデルが選ばれる。これは単に判別データを2群に1/0のダミー変数を与えると重回帰分析の回帰係数がLDFの判別係数になることを利用して重回帰分析の変数選択法の統計量を利用しているだけなのであくまで参考指標として用いる。これは「Fisherの仮説」を考えたことは判別関数が母集団と標本を考えた推測統計学的手法と考えることは間違いであり誤解であることと対応している。また判別手法には変数選択に用いる手法がないので、「一つとっておき法 (Leave one out, LOO)」が開発された。筆者はこれに代わって「小標本の100重交差検証法」を提案している。重回帰分析のCp統計量、AIC、BICも単に参考指標として用いる。LDF、QDF、LogiはFisherの線形判別関数、2次判別関数、ロジスティック回帰の誤分類数である。ロジスティック回帰は「総排気量」がモデルに入ってくると誤分類数が0と正しく判別し、それ以降も0と判別する。しかしLDFは5変数で初めて誤分類数が0になる。QDFは1変数と2変数で誤分類数が0になるのに、3変数で「定員」がモデルに入ってくると普通車の29人が小型車に誤判別される。

これはLDFもQDFもマハラノビスの汎距離による多群判別も以下にみるとおり、分散共散

散行列の逆行列を用いていて、ある変数が一定値の場合ランク落ちする。しかし、JMPには一般化逆行列の技術で対応し、さらに正則化判別分析という手法を用いている。多くのデータなデータでは成功してきたと思われるが、一方の群のある変数が一定値を取り他の群がばらつく場合に問題を生じるためである。このため、これらを使う品質管理のMT理論やゲノム判別も注意を必要とする。

SPSSの最初のLDFとLOOは事前確率を等確率に、次の「／比例」は29：15に比例させた誤分類数である。JMPでは両方ともSPSSの等確率と同じ結果になった。SPSSのQDFでは以前の版では問題があったが、本年4月以降では正しい結果を出している。筆者の提案する一定値をとる変数に小さなノイズを加えたのか、他のアルゴリズムを開発したかは不明である。

表5.1 判別結果

p	Var.	Cp	AICc	BIC	JMP			SPSS		SPSS / 比例		SPSS
					LDF	QDF	Logi	LDF	LOO	LDF	LOO	QDF
1	排気量	47.8	24.2	29	2	0	0	2	2	1	1	0
2	価格	14.9	4.3	10	1	0	0	1	2	0	2	0
3	定員	11.0	1.7	9	1	29	0	1	1	0	1	0
4	CO2	7.3	-1.4	7	1	29	0	1	1	0	1	0
5	燃費	5.0	-3.4	6	0	29	0	0	1	0	1	0
6	台数	7.0	-0.4	10	0	29	0	0	1	0	1	0

5.4. 主成分分析による布置

図5.2は元のデータを用いたスコアプロットである。ユーザー側に立って評価されるカムリとマーチが散布図の中に布置され、企業側に立って評価される3車種がどちらかというと周辺に布置される。これに対して図5.3は比を用いたスコアプロットである。5車種とも周辺に布置され、どちらかといえばこの方が表現としては適している。すなわち、全ての車種がこれらに包み込まれ、第2象限がユーザーに、第4象限が企業にとって利益を生む車種と考えられる。

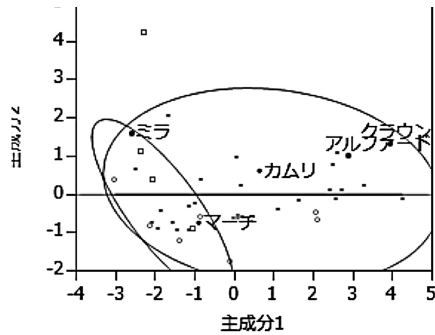


図5.2 元のデータによるスコアプロット

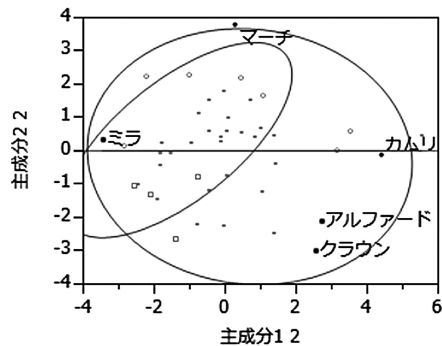


図5.3 比によるスコアプロット

6. まとめ

本章で用いたデータは、2011年の成蹊大学経済経営学研究科の後期開講科目「総合ビジネス分析（2単位）」で院生の岡野いつ佳さんが統計レポート作成に用いたデータである。統計ソフトのJMPを用いて、「学生の成績データ」という入門用のデータで統計手法を体系的に教えた。そして、自分で集めたデータで20頁以上の統計レポートの提出を求めている。提出されたレポートを添削し返却しているが、本人の今後の勉強のため、私が統計手法とDEA法の両方で再分析することとした。

販売台数と販売価格を目的変数とした重回帰分析が岡野さんのレポートの主要部分である。筆者は、DEA法を重回帰分析の対象であるデータ（重回帰型DEAデータ）にも適用すべきと考えている。さらに、普通車と小型車を2群判別することも考えられる。このような判

別型DEAデータは、優良企業と非優良企業の経営分析にも今後利用できると考えていた。以上のようにDEA法の適用分野の拡大を行うことで、DEAを教えた学生が企業で評価の可視化という改善運動で活躍できる場を作ろうと考えた。

しかし、セミナーで知遇を得た本田技研工業(株)の関連企業の、八千代工業(株)の辻井副社長(現、社長)に意見をいただいた。その結果、筆者の開発した「1入力固定改善法」を用いて、個々の性能の改善点を議論することは誤りであると考えるにいたった。製品の価格性能比を考える場合、CCRモデルと逆CCRモデルで、企業側に利益をもたらす車種と、ユーザーにとって価格の割に総合性能が優れている車種を対比して考えることにとどめ、統計手法との対比を行った。また、自動車よりも低価格なPCや携帯電話やスマホ、家電製品の方が適していると考えている。

(成蹊大学経済学部教授)

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), 2004). [新村秀一監修].
- [2] L.Schrage (2003), “*Optimizer Modeling with LINGO*”, LINDO Systems Inc..
- [3] S.Shinmura (2012), “Relationship between the DEA cluster and the lower limit of weights,” *Proceedings of DEA Symposium 2012*, 63-68.
- [4] K.Tone (1988), “Introduction to Efficiency Analysis of a company-DEA (1),” *Operations Research*, 32/12.
- [5] T.Ueda (2007), “Application of Multivariate analysis for DEA,” *Proceedings of DEA Symposium 2007*, 96-101.
- [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] _____ (2010), 『最適線形判別関数』. 日科技連出版社.
- [9] _____ (2011), 『数理計画法による問題解決法』. 日科技連出版社.
- [10] _____ (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.
- [11] _____ (2011), 『問題解決学としての統計入門. 第7回統計教育の方法論ワークショップ—問題解決力育成を目指した統計教育の方法論—』, 1-10.
- [12] _____ (2012), 『2011年と2012年度の統計入門の分析—2012年度の半肢授業の影響分析からセンター試験の分析へ—』. 第8回統計教育の方法論ワークショップ—問題解決

力育成を目指した統計教育の方法論—], 36-45.

- [13] _____ (2013), 「DEA利用のための実践的解説書—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』
- [14] _____ (2012), 「Fisherの判別分析を超えて (2) —大学入試センター 105試験データの合否判定—」. 2012年OR学会春春季発表会予稿集.
- [15] 竹内啓 (2011), 統計. 書評, 71-74.
- [16] Friedman, J. H. (1989), “Regularized Discriminant Analysis,” *Journal of the American Statistical Association*, 84/405, 165-175.
- [17] 新村秀一 (2007), 『JMPによる統計レポート作成法』. 丸善.
- [18] 新村秀一 (2012), 「SAS/JMPとの歩み」. *SAS Technical Report* (春号), 13-16.