

大型小売店における売上予測 (第4報)

Sales Forecast in a Large Size Retail store (the 4th report)

橋本郁郎

Ikuro HASHIMOTO

Abstract A Sales forecasting plays important roles for making a management plan and proceeding to research development and product planing. There are various forecasts such as technical forecast, economic forecast, sales forecast and product forecast. Succeeding to the previous report, the present study discusses further sales forecast of department stores in Aichi Prefecture by means of a multiple regression analysis. The sales forecast for 1987 was discussed in the previous report and the present study deals with that for 1991. About this year of 1991, economy started folling down. As the results of the analysis, the range of the forecast errors shows at max. of 30.972% and min. of 4.640%. It seems that the maximem error rate sounds to be too lorge. The improvement of decreasing forecast errors by adding a dummy variable, effects of climate and others is discussed in the present paper.

1. はじめに

現代社会における企業にとって、経営計画を立てる場合や研究開発活動をするのに、予測は非常に重要なものとなっている。第1報から第3報において^{1) 2) 3)}愛知県内の百貨店の衣料品の総売上につき、1979年～1987年の間の検討を行った。今回は、愛知県内の百貨店の総売上につき、バブル崩壊の影響が現れ始めた1979年～1991年の間につき研究を行った。

2. 研究方法

2・1 方法

1979年より1990年までの12年間の月次データにより、1991年の愛知県内の百貨店の総売上高を予測す

ることである。経済時系列の分析に於いて、原系列は一般には傾向変動 (T : Trend), 循環変動 (C : Cycle movement), 季節変動 (S : Seasonal movement), 不規則変動 (I : Irregular component) を含んでいる。

このTCSI系列の解析には、第1報に示す予測手順に従って行った。

2・2 季節調整

本研究は百貨店の総売上高を予測するのが目的であるので、月次データは当然季節変動の影響を受ける。季節変動の影響を除去するには多くの方法があるが、その中で対移動平均比率法を用い

表1 季節調整指数

1	0.893
2	0.786
3	1.043
4	0.900
5	0.889
6	0.870
7	1.332
8	0.736
9	0.835
10	0.939
11	0.952
12	1.826

ることにする。季節指数を表1に、売上高の原データと季節済みデータを図1に示す。

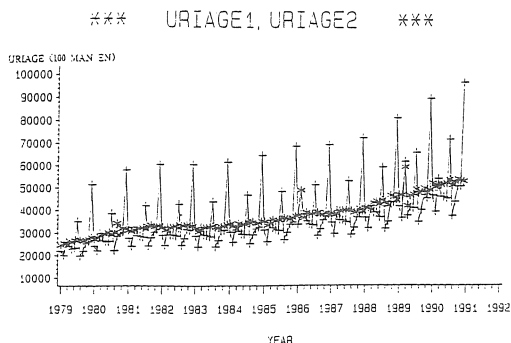


図1 売上高の原データと季節済みデータ

3. 回帰モデル式の構築

3.1 回帰モデル

季節済み系列(TCI)には、単なる傾向変動の外にも影響を与える種々の要因があると考えられるので、これを解明するため以下の如く回帰モデルを導入する。

$$y_i = \alpha_0 + \alpha_1 X_{i1} + \dots + \alpha_p X_{ip} + \varepsilon_i \dots (1)$$

(i = 1, 2, \dots, n)

このn組のデータによって求められた $\alpha_0, \alpha_1, \dots, \alpha_p$ の最小二乗指定値を a_0, a_1, \dots, a_p で表すと、yの X_1, X_2, \dots, X_p に対する重回帰式が(2)式の如く得られる。

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_p X_p \dots (2)$$

以下の計算はSASによって実施した。

3.2 最適説明変数

3.2.1 最適説明変数の導入

売上予測をするための重回帰分析を行うのに必要な説明変数として次の5つを取り上げることにした。

- | | |
|-----------------|---------|
| 1) 可処分所得(季節済み) | SYOTOK2 |
| 2) 消費支出(季節済み) | SHISYU2 |
| 3) 売場面積 | MENSEKI |
| 4) 消費者物価指数(愛知県) | SHISU |
| 5) 時間 | TIME |

また被説明変数としては、季節済みの売上高(URIMAGE2)を用いた。

3.2.2 説明変数の散布図

各説明変数の時間に対する散布図を図2~5に示す。

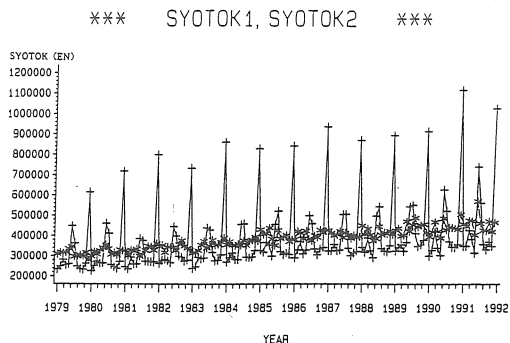


図2 可処分所得の原データと季節済みデータ

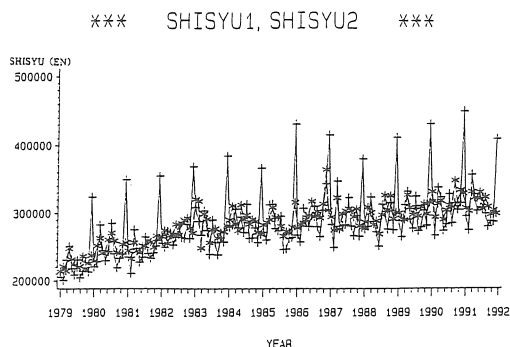


図3 消費支出の原データと季節済みデータ

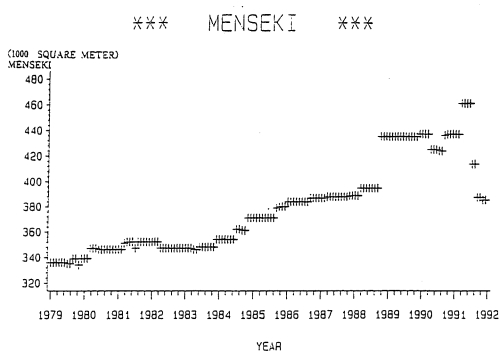


図4 売場面積

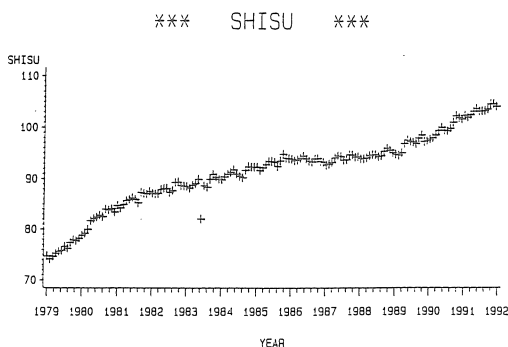


図5 消費者物価指数

3・2・3 基本統計量と相関係数

被説明変数と説明変数の基本統計量及び相関係数を表2に示す。

表2 基本統計量及び相関係数

SIMPLE STATISTICS						
VARIABLE	N	MEAN	STD DEV	SUM	MINIMUM	MAXIMUM
URIAGE2	144	37235	7596.822615	5361849	24507	59262
SYOTOK2	156	380937	52004	59426186	277570	570616
SHISYU2	156	283747	29795	44264571	213431	363467
MENSEKI	156	379.217949	35.470714	59158	334.000000	461.000000
TIME	156	78.500000	45.177428	12246	1.000000	156.000000
SHISU	156	91.400000	7.149807	14258	74.100000	104.700000

PEARSON CORRELATION COEFFICIENTS / PROB > R UNDER H0: RHO=0 / NUMBER OF OBSERVATIONS						
	URIAGE2	SYOTOK2	SHISYU2	MENSEKI	TIME	SHISU
URIAGE2	1.00000 0.0 144	0.82167 0.0001 144	0.73266 0.0001 144	0.95148 0.0001 144	0.93604 0.0001 144	0.87427 0.0001 144
SYOTOK2	0.82167 0.0001 144	1.00000 0.0 156	0.75161 0.0001 156	0.83934 0.0001 156	0.88624 0.0001 156	0.85137 0.0001 156
SHISYU2	0.73266 0.0001 144	0.75161 0.0001 156	1.00000 0.0 156	0.73816 0.0001 156	0.82708 0.0001 156	0.84846 0.0001 156
MENSEKI	0.95148 0.0001 144	0.83934 0.0001 156	0.73816 0.0001 156	1.00000 0.0 156	0.91295 0.0001 156	0.83391 0.0001 156
TIME	0.93604 0.0001 144	0.88624 0.0001 156	0.82708 0.0001 156	0.91295 0.0001 156	1.00000 0.0 156	0.95695 0.0001 156
SHISU	0.87427 0.0001 144	0.85137 0.0001 156	0.84846 0.0001 156	0.83391 0.0001 156	0.95695 0.0001 156	1.00000 0.0 156

3・2・4 寄与率

説明変数の全ての組み合わせに対する寄与率を表3に示す。寄与率が1に近い程その回帰モデルはよく説明されていることになるが、これのみで全てを判定することは出来ない。

表3 説明変数の組合せに対する寄与率

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
1	0.90531852	MENSEKI
1	0.87617370	TIME
1	0.74434059	SHISU
1	0.67513612	SYOTOK2
1	0.53679779	SHISYU2
2	0.92389917	MENSEKI SHISU
2	0.92087251	MENSEKI TIME
2	0.91077026	SHISYU2 MENSEKI
2	0.90666361	SYOTOK2 MENSEKI
2	0.88040102	SHISYU2 TIME
2	0.87718528	TIME SHISU
2	0.87670299	SYOTOK2 TIME
2	0.78836253	SYOTOK2 SHISU
2	0.76495321	SHISYU2 SHISU
2	0.70867958	SYOTOK2 SHISYU2
3	0.92483227	SYOTOK2 MENSEKI SHISU
3	0.92439525	SHISYU2 MENSEKI SHISU
3	0.92430616	MENSEKI TIME SHISU
3	0.92204992	SYOTOK2 MENSEKI TIME
3	0.92089861	SHISYU2 MENSEKI TIME
3	0.91087184	SYOTOK2 SHISYU2 MENSEKI
3	0.88080934	SYOTOK2 SHISYU2 TIME
3	0.88043519	SHISYU2 TIME SHISU
3	0.87766927	SYOTOK2 TIME SHISU
3	0.78972311	SYOTOK2 SHISYU2 SHISU
4	0.92573586	SYOTOK2 MENSEKI TIME SHISU
4	0.92521615	SYOTOK2 SHISYU2 MENSEKI SHISU
4	0.9208853	SHISYU2 MENSEKI TIME SHISU
4	0.92205903	SYOTOK2 SHISYU2 MENSEKI TIME
4	0.88083991	SYOTOK2 SHISYU2 TIME SHISU
5	0.92633493	SYOTOK2 SHISYU2 MENSEKI TIME SHISU

3・3 最適モデル式

5個の説明変数に次に述べる3条件をつけて検討した。

- 1) 売場面積は必ず取り入れる
- 2) 可処分所得と消費支出はどちらか一方
- 3) 時間と消費者物価指数はどちらか一方

説明変数3個でこの条件を満たすものにつき重回帰分析を行い、その結果を表4の(a)(b)(c)(d)に示す。これによるとR-SQUAREとDURBIN-WATSONでは(d)が一番優れているが、各パラメータのPROB>|T|では、

(b) が一番小さくて良い値を示しているの、(b) と

T C 系列の予測値を図 6 に示す。T C 系列の予測値に季節指数を掛けたものを売上予測値とし(4)式により予測誤差を算出する。

表 4 説明変数の組合せに対する重回帰分析

PARAMETER ESTIMATES							
VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > T	STANDARDIZED ESTIMATE	
(a) INTERCEP	1	-19373	6592.6243396	-2.939	0.0039	0.00000000	
SHISYU2	1	-0.002413	0.0123029	-0.215	0.8302	-0.0036235	
MENSEKI	1	139.463615	16.47306294	8.466	0.0001	0.60803735	
TIME	1	67.945859	16.04802759	4.234	0.0001	0.37308315	
DURBIN-HATSON D		1.439					
PARAMETER ESTIMATES							
VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > T	STANDARDIZED ESTIMATE	
(b) INTERCEP	1	-17071	5510.9686995	-3.098	0.0024	0.00000000	
SYOTOK2	1	-0.011960	0.00821071	-1.454	0.1481	-0.07506127	
MENSEKI	1	161.686034	15.49991779	9.025	0.0001	0.61772670	
TIME	1	77.041322	14.65550359	5.257	0.0001	0.42302532	
DURBIN-HATSON D		1.500					
PARAMETER ESTIMATES							
VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > T	STANDARDIZED ESTIMATE	
(c) INTERCEP	1	-53071	2591.7323753	-20.784	0.0001	0.00000000	
SHISYU2	1	-0.010995	0.01167174	-0.958	0.3395	-0.04265203	
MENSEKI	1	169.499706	9.86427562	17.183	0.0001	0.73896771	
SHISU	1	338.105377	67.31231637	5.023	0.0001	0.28886834	
DURBIN-HATSON D		1.560					
PARAMETER ESTIMATES							
VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > T	STANDARDIZED ESTIMATE	
(d) INTERCEP	1	-54043	2825.2687867	-19.447	0.0001	0.00000000	
SYOTOK2	1	-0.010211	0.00774579	-1.318	0.1894	-0.06419376	
MENSEKI	1	176.198064	11.05184206	15.943	0.0001	0.76819328	
SHISU	1	330.240482	96.77025762	3.417	0.0001	0.28212942	
DURBIN-HATSON D		1.602					

(d) を取り上げ解析を進めることにする。後述(4)式により予測誤差を計算すると(b)の方が最大誤差が小さいことが判明したので、以下の(b)につき検討を進める。回帰式は(3)の如くなる。

$$Y = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + \dots \quad (3)$$

X_1 : SYOTOK2 X_2 : MENSEKI
 X_3 : TIME Y : URIAGE2

この計算結果を表 5 に示す。

表 5 最適モデル式の重回帰分析結果

ANALYSIS OF VARIANCE							
SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F		
MODEL	3	7609470601.6	2536490200.5	552.007	0.0001		
ERROR	140	643304478.8	4595031.9914				
C TOTAL	143	8252775080.4					
ROOT MSE		2143.60257		R-SQUARE 0.9220			
DEP MEAN		37235.06250		ADJ R-SQ 0.9204			
C.V.		5.75695					
PARAMETER ESTIMATES							
VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR HO: PARAMETER=0	PROB > T	STANDARDIZED ESTIMATE	
INTERCEP	1	-17071	5510.9686995	-3.098	0.0024	0.00000000	
SYOTOK2	1	-0.011940	0.00821071	-1.454	0.1481	-0.07506127	
MENSEKI	1	161.686034	15.49991779	9.025	0.0001	0.61772670	
TIME	1	77.041322	14.65550359	5.257	0.0001	0.42302532	
DURBIN-HATSON D		1.500					
(FOR NUMBER OF OBS.)		144					
1ST ORDER AUTOCORRELATION		0.237					

4. 予測結果

1991年度の売上予測をするには本来は説明変数も予測値を用いるのであるが、本論の場合は予測式適度を見るために実績値を用いることにした。

*** URIAGE, PREDICT ***

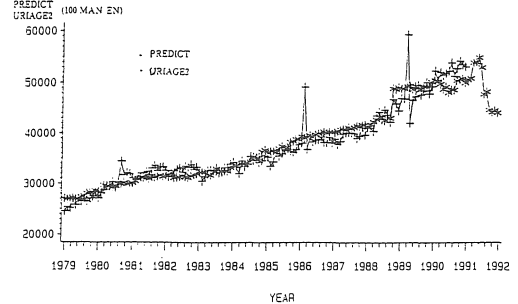


図 6 予測値及び実績値(季節済み)

$$\text{予測誤差} = \frac{\text{実績値} - \text{予測値}}{\text{予測値}} \times 100 \quad \dots (4)$$

(予測値 = T C の補外値 × 季節調整指数)

また、予測誤差を表 6 に示す。

表 6 予測誤差

GOSA	
1	11.2682
2	6.1621
3	6.8686
4	4.6403
5	4.6978
6	9.7149
7	19.3700
8	16.9963
9	25.2067
10	30.9721
11	27.1328
12	21.4575

5. 残差の検討

重回帰分析において、当てはめられた回帰式の正当性を見るうえで、残差の検討は有用である。

5.1 残差の時系列プロット

時間を横軸にとって、そのときの残差の値を打点すると、残差の時間的変化を知ることが出来る。

これを図 7 に示す。残差は 0 を中心にして上下に変動するが、これらについて以下検討する。

(1) 傾向的な変化

右上がり、あるいは右下がりの傾向、周期的変化、時間と共に増大減少の傾向などは認められないが、

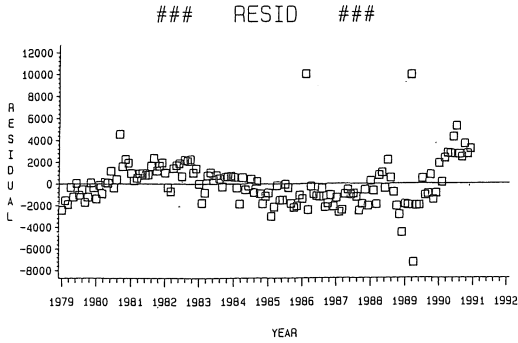


図7 残差の時系列プロット

僅かに曲線の傾向が認められる。

(2) 符号

+符号 n_1 が67個、-符号 n_2 が77個となっている。
この符号は(5)式で検定を行うものとする。

$$P = \frac{n - 1}{2} - K\sqrt{n + I} \quad \dots\dots(5)$$

$$n = n_1 + n_2$$

有意水準	0.01	0.05
K	1.2879	0.9800

よって有意水準1%では

$$P = \frac{144-1}{2} - 1.2879\sqrt{144+I} = 55.992$$

5%では

$$P = \frac{144-1}{2} - 0.9800\sqrt{144+I} = 59.699$$

すなわち n_1, n_2 の小さい方の数値が、有意水準1%では55以下、有意水準5%では59以下ならばそれぞれ有意となり、50:50とは認められない。本論の場合は小さい方の数値は $n_1 = 67 > 55$ となり1%の有意水準で有意ではなくなり、+-は50:50と認められる。

(3) 連の数による検定

+の連、-の連を合わせた全体の連の数を n_R で表すと、+、-がランダムにでる場合は n_R は1つの確率変数になり、ある分布に従う。

$n_1 > 10, n_2 > 10$ の場合は連の数の分布は近似的に(6)式の正規分布に従う。

$$\mu = \frac{2 n_1 n_2}{n_1 + n_2} + 1$$

$$\sigma^2 = \frac{2 n_1 n_2 (2 n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)} \quad \dots\dots(6)$$

従って(7)式の如く u は標準正規分布に従う。

$$u = \frac{|n_R - \mu| - 0.5}{\sigma} \quad \dots\dots(7)$$

これに $n_1 = 67, n_2 = 77, n_R = 36$ を代入すれば、 $\mu = 72.65, \sigma^2 = 35.40, u = 6.075$ となり、 u は標準正規分布の両側1%の点である2.576より大きくなるので有意である。即ち連の数から見ると、この残差系列はランダムとは見なせない。

(4) 連の長さによる検定

+、-の符号がランダムな系列では、一方の符号だけが連続して多く現れることはまれである。7個以上も同じ符号がならぶ場合はランダムとは見なせないとする、本論の場合は+側で17個、-側で22個が連続して現れているので、連の長さからは大いに問題を含んでいると考えられる。

5・2 残差の異状値

表5(b)の分散分析表より、誤差の不偏分散 V_e は

$$V_e = 4595031.9914$$

$$\sqrt{V_e} = 2143.6025$$

が得られるから、残差の3シグマ限界を求めると

$$\pm 3\sqrt{V_e} = \pm 6430.8077$$

でこれを越えるものは

$$\text{No. 86} \quad +10051.64$$

$$\text{No. 123} \quad +9946.13$$

$$\text{No. 124} \quad -7284.82$$

の3個で、これらは明らかに異状値である。

5・3 ダービンワトソン比

残差の連なりがランダムかどうかをコンピュータで検定するにはダービン・ワトソン比(D)を用いる。残差が全くランダムであれば、Dの値は2.0となり、正の相関があれば2よりも小さくなり、負の相関があれば2よりも大きくなる。本論の場合はDの値は1.500となり、統計限界は約

$$d_L = 1.48(1\%) \quad 1.61(5\%)$$

$$d_U = 1.60(1\%) \quad 1.74(5\%)$$

となる。したがって1%の統計限界では

$$d_L < D < d_U$$

となり、いづれとも判定出来ないが、5%の統計限界では

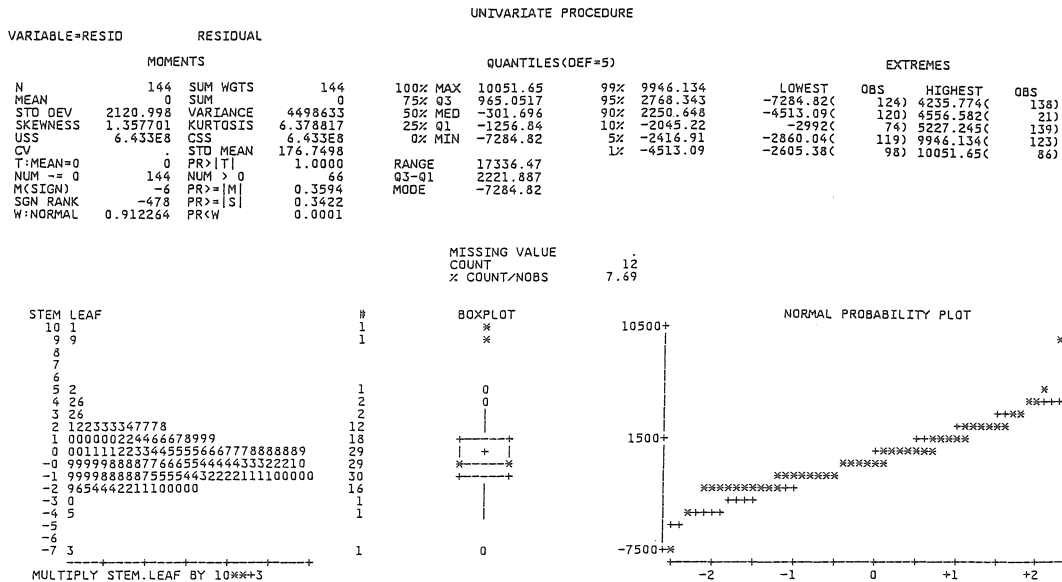
$$d_L > D$$

となり、正の自己相関が認められる。

5・4 残差の各種統計値

SASのUNIVARIATE PROCEDUREを用いて、残差の各種統計量を計算したものを表7に示す。

表7 残差の各種統計量



6. おわりに

企業が経営計画を立てる場合には、予測が重要なものとなっている。百貨店の売上予測する場合に、説明変数として可処分所得、売場面積、時間を用い1979年より1990年までの12年間のデータを用い1991年の売上を予測した。その結果予測誤差は、最小は4月の+4.64%から最大は10月の+30.97%であることが判明した。

今後の課題としては、残差分析の結果正の自己相関が認められ、モデル式に多少問題があると認められる。またバブル崩壊の影響があったとは云え、予測誤差も大きすぎる。これらについては、ダミー変数の導入や、気温の影響を考慮することなどにより改善をはかるつもりである。

参考文献

1) 橋本郁郎：大型小売店における売上予測，愛

知工業大学研究報告 Vol. 25, Part B, P. 71~77, 1990

2) 橋本郁郎：大型小売店における売上予測（第2報），愛知工業大学研究報告 Vol. 26, Part B, P. 61~67, 1991

3) 橋本郁郎：大型小売店における売上予測（第3報），愛知工業大学研究報告 Vol. 27, Part B, P. 93~98, 1992

4) 大西正和：需要予測とコンピュータプログラム，日刊工業新聞，東京，1982

5) 通産省産業政策局：商業動態統計年報，通産産業調査会，東京，1979~1991

6) SAS STAT USER'S GUIDE Ver. 6.03, SAS Institute Inc.

7) SAS GRAPH REFERENCE GUIDE Ver. 6.03 SAS Institute Inc.

8) 奥野忠一，久米均，芳賀敏郎，吉澤正：多変量解析法，日科技連，東京，1986

(受理 平成7年3月20日)