大型小売店における売上予測

橋　本　郁　郎

Sales Forecast in a Large Size Retail Store

Ikuro HASHIMOTO

A sales forecasting is very important for making a management plan and proceeding to research & development. There are various forecasts such as technical forecast, economic forecast, sales forecast and production forecast. The present study discusses sales forecast of a clothing in a department store in Aichi Prefecture by means of a multiple regression analysis. As the results of the analysis, forecast errors obtained, are max. 6.522% and min. −0.05%.

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，予測が非常に重要なものとなっている。予測には技術予測，経済予測，売上予測，生産予測等いろいろあり，それぞれの目的に合せて用いられる。今回はその中で売上予測をとり上げ，具体的に大型小売店のうち愛知県内の百貨店の衣料品の総売上予測につき研究を行なった。

２．研究方法

２・１　方法

昭和54年より61年までの月次データにより，昭和62年の愛知県内の百貨店の衣料品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。よって図１に示す手順に従って解析を進めることとした[1]。

２・２　季節調整

本研究は百貨店の衣料品の売上高を予測するのが目的であるので，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のには多くの方法が提案されているが，その中で対移動平均比率法を用いることにする[2]。売上高の原データ[3]と季調済みデータを図２に，また季節指数を表１に示す。

３．回帰モデル式の構築

３・１　回帰モデル

季調済み系列（ＴＣＩ）には，単なる傾向変動の

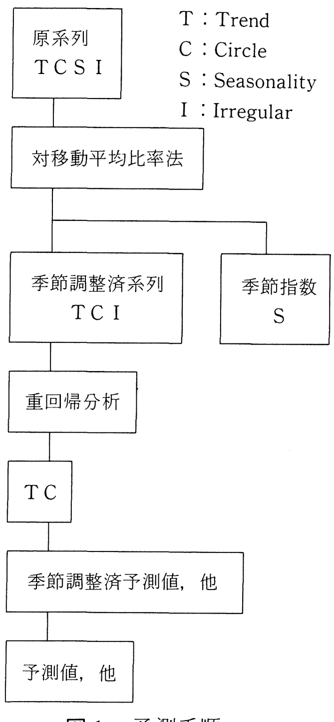

図１　予測手順

経営工学科

表１　季節調整指数

１　月	２　月	３　月	４　月
1.003	0.758	1.097	0.952
５　月	６　月	７　月	８　月
0.940	0.908	1.106	0.665
９　月	10　月	11　月	12　月
0.883	1.039	1.041	1.607

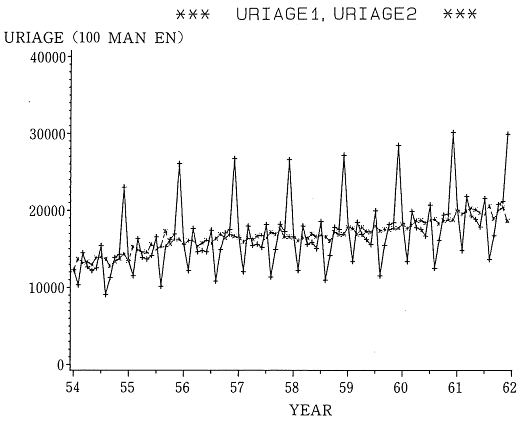

図２　売上高の原データと季調済みデータ

外にもこれに影響を与える各種の要因があると考えられる。これらを解明するため以下の如く回帰モデル式を導入する。

被説明変数ｙと説明変数 x_1, x_2, ……x_p との間に(1)式の如き線形関係が成立すると仮定する。

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \cdots\cdots + \alpha_p x_{ip} + \varepsilon_i \cdots\cdots (1)$$

$$(i = 1,\ 2,\ \cdots\cdots\cdots n)$$

このｎ組のデータによって求められた α_0, α_1, ………α_p の最小二乗推定値を a_0, a_1, ………a_p で表すと，ｙの x_1, x_2, …………x_p に対する重回帰式が(2)式の如く得られる。

$$Y = a_0 + a_1 x_1 + a_2 x_2 + \cdots\cdots\cdots + a_p x_p \cdots\cdots\cdots\cdots (2)$$

以下の統計計算はＳＡＳ[4)5)]（Statistical Analysis System）によって行なった。

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なうのに必要な説明変数として次の５つを考えた。

１）可処分所得（季調済み）	SYOTOK2
２）消費支出（季調済み）	SHISYU2
３）売場面積	MENSEKI
４）消費者物価指数	SHISU
５）時間	TIME

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３～６に示す[3)]。

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。

３・２・４　寄与率

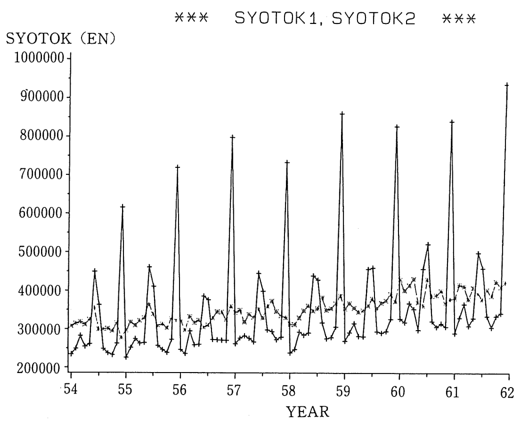

図３　可処分所得の原データと季調済みデータ

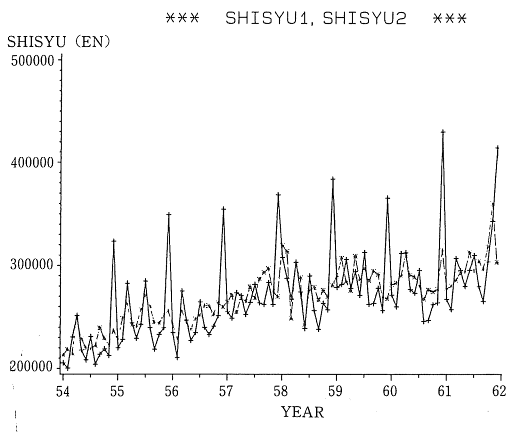

図４　消費支出の原データと季調済みデータ

説明変数の全ての組合せに対する寄与率を表３に示す。寄与率が大きいほど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

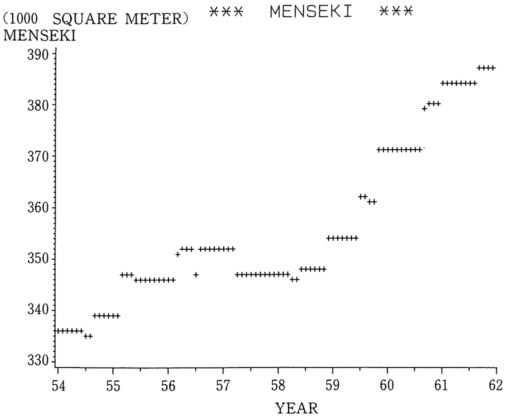

図 5　売場面積

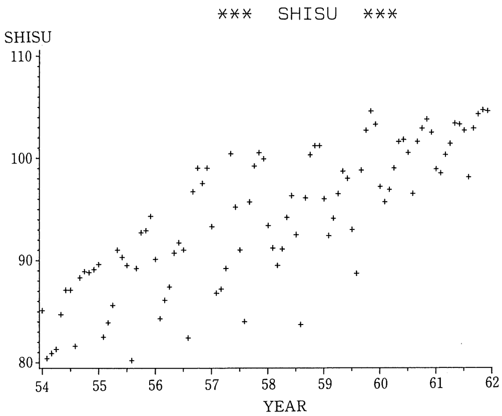

図 6　消費者物価指数（昭和60年＝100）

することは出来ない。

3・3　最適モデル式

表 3 に説明変数の全ての組合せについてその寄与率を示した。このうち説明変数 4 個と 5 個の組合せにつき，重回帰分析を行ない，寄与率以外のパラメータを検討し，最適モデルを決定するものとする。重回帰分析の結果を表 4 に示す。

このうち回帰係数の符号が経済理論と矛盾するものとして

(3)　SYOTOK2　−0.000906204

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	0.60613351	SHISU
1	0.63491264	SHISYU2
1	0.65487497	SYOTOK2
1	0.80571584	MENSEKI
1	0.88098418	TIME
2	0.74552889	SYOTOK2 SHISU
2	0.75504993	SHISYU2 SHISU
2	0.76491636	SYOTOK2 SHISYU2
2	0.82092116	SYOTOK2 MENSEKI
2	0.83237617	MENSEKI SHISU
2	0.87340915	SHISYU2 MENSEKI
2	0.88127248	SYOTOK2 TIME
2	0.88170303	TIME SHISU
2	0.88347633	SHISYU2 TIME
2	0.89347376	MENSEKI TIME
3	0.80718421	SYOTOK2 SHISYU2 SHISU
3	0.84064018	SYOTOK2 MENSEKI SHISU
3	0.87531271	SYOTOK2 SHISYU2 MENSEKI
3	0.88190313	SHISYU2 MENSEKI SHISU
3	0.88198782	SYOTOK2 TIME SHISU
3	0.88383196	SYOTOK2 SHISYU2 TIME
3	0.88437396	SHISYU2 TIME SHISU
3	0.89355674	SYOTOK2 MENSEKI TIME
3	0.89397271	MENSEKI TIME SHISU
3	0.90061042	SHISYU2 MENSEKI TIME
4	0.88290919	SYOTOK2 SHISYU2 MENSEKI SHISU
4	0.88472781	SYOTOK2 SHISYU2 TIME SHISU
4	0.89405344	SYOTOK2 MENSEKI TIME SHISU
4	0.90074435	SYOTOK2 SHISYU2 MENSEKI TIME
4	0.90133074	SHISYU2 MENSEKI TIME SHISU
5	0.90146217	SYOTOK2 SHISYU2 MENSEKI TIME SHISU

表 2　基本統計量及び相関係数

VARIABLE	N	MEAN	STD DEV	SUM	MINIMUM	MAXIMUM
URIAGE2	96	16687.2708333	1880.59484699	1601978.00000	12239.0000000	20465.0000000
SYOTOK2	108	356417.2962963	38209.37272824	38493068.00000	276819.0000000	430743.0000000
SHISYU2	108	272547.4259259	27891.20989498	29435122.00000	213117.0000000	360157.0000000
MENSEKI	108	359.4259259	17.77176365	38818.00000	335.0000000	389.0000000
TIME	108	54.5000000	31.32091953	5886.00000	1.0000000	108.0000000
SHISU	108	95.0462963	7.16020187	10265.00000	80.2000000	106.5000000

PEARSON CORRELATION COEFFICIENTS / PROB > |R| UNDER H0:RHO=0 / NUMBER OF OBSERVATIONS

	URIAGE2	SYOTOK2	SHISYU2	MENSEKI	TIME	SHISU
URIAGE2	1.00000	0.80924	0.79681	0.89762	0.93861	0.77855
	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001
	96	96	96	96	96	96
SYOTOK2	0.80924	1.00000	0.70481	0.84675	0.86879	0.73103
	0.0001	0.0000	0.0001	0.0001	0.0001	0.0001
	96	108	108	108	108	108
SHISYU2	0.79681	0.70481	1.00000	0.68160	0.79923	0.66320
	0.0001	0.0001	0.0000	0.0001	0.0001	0.0001
	96	108	108	108	108	108
MENSEKI	0.89762	0.84675	0.68160	1.00000	0.93436	0.79080
	0.0001	0.0001	0.0001	0.0000	0.0001	0.0001
	96	108	108	108	108	108
TIME	0.93861	0.86879	0.79923	0.93436	1.00000	0.84368
	0.0001	0.0001	0.0001	0.0001	0.0000	0.0001
	96	108	108	108	108	108
SHISU	0.77855	0.73103	0.66320	0.79080	0.84368	1.00000
	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
	96	108	108	108	108	108

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 H0: PARAMETER=0	PROB > |T|	STANDARDIZED ESTIMATE	VARIANCE INFLATION
(1)	INTERCEP	1	-16884.48012	1921.41259	-8.788	0.0001	0	0
	SYOTOK2	1	0.003046959	0.003445830	0.884	0.3789	0.05944749	3.51269297
	SHISYU2	1	0.02036553	0.003553247	5.732	0.0001	0.30176689	2.15437376
	MENSEKI	1	65.89949264	8.59019809	7.671	0.0001	0.54248417	3.88628778
	SHISU	1	37.85120681	15.57808460	2.430	0.0171	0.13795814	2.50542775
(2)	INTERCEP	1	10457.96362	2064.86153	5.065	0.0001	0	0
	SYOTOK2	1	0.001847060	0.003494688	0.529	0.5984	0.03603694	3.67001363
	SHISYU2	1	0.006163028	0.004190455	1.471	0.1448	0.09132084	3.04362259
	TIME	1	53.42005980	6.82767678	7.824	0.0001	0.79129786	8.07484151
	SHISU	1	14.12504100	16.79622489	0.841	0.4026	0.05148222	2.95852602
(3)	INTERCEP	1	2254.00775	3479.56896	0.648	0.5188	0	0
	SYOTOK2	1	-0.000906204	0.003441321	-0.263	0.7929	-0.01768043	3.87203053
	MENSEKI	1	32.41343310	10.06869204	3.219	0.0018	0.26682716	5.90079649
	TIME	1	45.96305423	6.78589428	6.773	0.0001	0.68083912	8.67840549
	SHISU	1	10.50602311	16.08481227	0.653	0.5153	0.03829181	2.95203623
(4)	INTERCEP	1	-1657.40432	3647.34129	-0.454	0.6506	0	0
	SYOTOK2	1	-0.001167675	0.003332242	-0.350	0.7268	-0.02278185	3.87519200
	SHISYU2	1	0.01037241	0.004040587	2.567	0.0119	0.15369348	3.28644648
	MENSEKI	1	39.92064932	10.13799283	3.938	0.0002	0.32862651	6.38557648
	TIME	1	36.08730063	7.47350158	4.829	0.0001	0.53455207	11.23584274
(5)	INTERCEP	1	-2767.51003	3851.18157	-0.719	0.4742	0	0
	SHISYU2	1	0.01050489	0.004032561	2.605	0.0107	0.15565652	3.29285822
	MENSEKI	1	38.91146459	9.83957035	3.955	0.0002	0.32031891	6.05092506
	TIME	1	32.64154622	7.71136567	4.233	0.0001	0.48351098	12.03353814
	SHISU	1	12.66940288	15.54404215	0.815	0.4172	0.04617679	2.96021057
(6)	INTERCEP	1	-2700.64630	3874.74666	-0.697	0.4876	0	0
	SYOTOK2	1	-0.001156737	0.003338593	-0.346	0.7298	-0.02256843	3.87525545
	SHISYU2	1	0.01054541	0.004053888	2.601	0.0109	0.15625691	3.29560077
	MENSEKI	1	39.72159742	10.16020480	3.910	0.0002	0.32698792	6.38931696
	TIME	1	33.50409494	8.13900598	4.116	0.0001	0.49628770	13.27557936
	SHISU	1	12.64750422	15.61987371	0.810	0.4202	0.04609697	2.96025904

表5　主要なパラメータ

	(1)	(2)	(5)
R-SQUARE	0.8829	0.8847	0.9013
DURBIN-WATSON	1.312	1.021	1.247

(4)　SYOTOK2　−0.001167675

(6)　SYOTOK2　−0.001156737

は検討から除かれる。それぞれの(1)(2)(5)の主要なパラメータを表5に示す。

R-SQUAREは(5)が0.9013で最も良いが，DURBIN—WATSONは(1)が1.312で最も良い。しかし(5)も1.247を示しており，全体として(5)を採用することにする。

故に最適モデル式は(3)式の如くなる。

$$Y=a_0+a_1x_1+a_2x_2+a_3x_3+a_4x_4 \quad\cdots\cdots\cdots\cdots\cdots(3)$$

x_1：SHISYU2　　x_2：MENSEKI

x_3：TIME　　x_4：SHISU

Y：URIAGE2

この計算結果を表6に示す。

4．予測結果

昭和62年度の売上を予測するには，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，予測式の適合度をみるために実績値を用いることとした。TC系列の予測値を図7に示す。このTC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 6　最適モデルの重回帰分析結果

SAS

DEP VARIABLE: URIAGE2

ANALYSIS OF VARIANCE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	4	302829563	75707390.79	207.818	0.0001
ERROR	91	33150949.81	364296.15		
C TOTAL	95	335980513			

ROOT MSE	603.5695	R-SQUARE	0.9013
DEP MEAN	16687.27	ADJ R-SQ	0.8970
C.V.	3.616946		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|	STANDARDIZED ESTIMATE	VARIANCE INFLATION
INTERCEP	1	-2767.51003	3851.18157	-0.719	0.4742	0	0
SHISYU2	1	0.01050489	0.004032561	2.605	0.0107	0.15565652	3.29285822
MENSEKI	1	38.91146459	9.83957035	3.955	0.0002	0.32031891	6.05092506
TIME	1	32.64154622	7.71136567	4.233	0.0001	0.48351098	12.03353814
SHISU	1	12.66940288	15.54404215	0.815	0.4172	0.04617679	2.96021057

DURBIN-WATSON D	1.247
(FOR NUMBER OF OBS.)	96
1ST ORDER AUTOCORRELATION	0.317

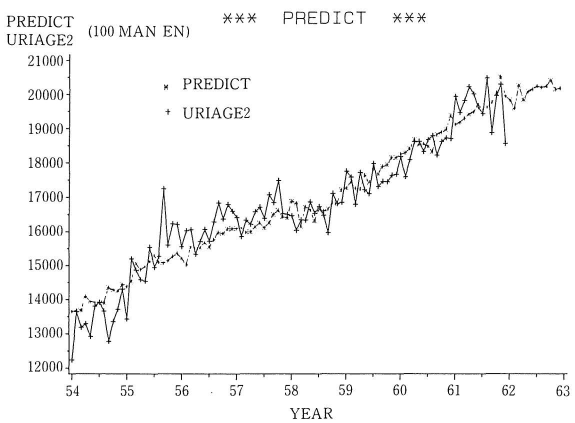

図 7　予測値及び実績値（季調済み）

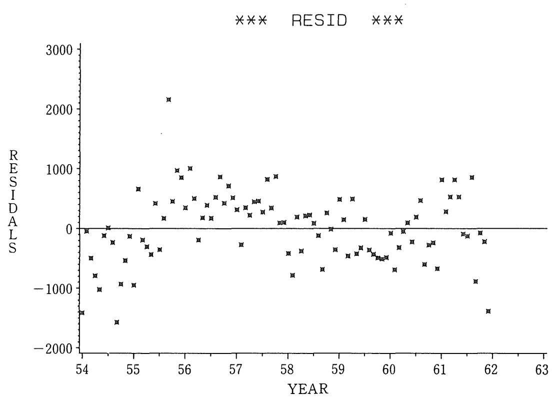

図 8　残差の時系列プロット

表 7　予測誤差（%）

1　月	2　月	3　月	4　月
3.99	4.47	1.32	6.52
5　月	6　月	7　月	8　月
6.70	2.52	−0.05	2.32
9　月	10　月	11　月	12　月
−1.77	2.81	5.86	−2.93

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

（予測値＝TCの補外値×季節調整指数）

また，予測誤差を表 7 に示す。

5．残差の検討[6)]

重回帰分析において残差の検討は、当てはめた回帰式の正当性をみるうえで重要である。

5・1　残差の時系列プロット

時間を横軸にとって残差の値をプロットすると、時間的変化を知ることが出来る。これを図 8 に示す。

(1)　傾向的な変化

右上り右下りの傾向，周期的変化，時間と共に増大減少の傾向などは，いずれも認められない。

(2)　符号

＋符号 n_1が48，－符号 n_2が48で頂度50：50になっており，この点での問題はない。

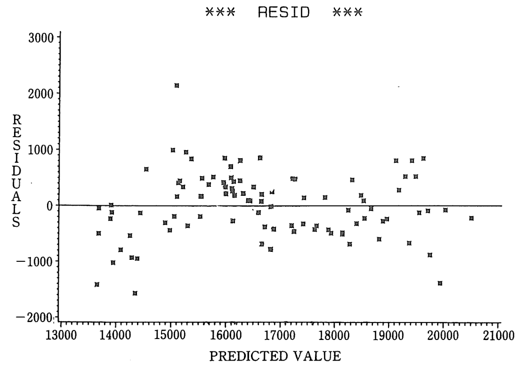

図 9　残差と予測値の散布図

(3)　連の数による検定

＋の連，－の連を合せた全体の連の数を n_R で表わすと，＋，－がランダムに出る場合は 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(5)式の正規分布に従う。

$$\left.\begin{aligned}\mu&=\frac{2n_1n_2}{n_1+n_2}+1\\ \sigma^2&=\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}\end{aligned}\right\}\cdots\cdots\cdots\cdots(5)$$

従って(6)式の如く u は標準正規分布にしたがう。

$$u=\frac{|n_R-\mu|-0.5}{\sigma}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(6)$$

これに，$n_1=48$，$n_2=48$，$n_R=33$ を代入すれば，$\mu=49$，$\sigma^2=23.747$，$u=3.1807$ 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。即ち連の数からみると，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わ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さないとすると，本論の場合は＋側で８個，10個，－側で９個が連続して現われているので，連の長さからも多少問題を含んでいる。

５・２　残差と従属変数の値との散布図

横軸に従属変数の予測値，縦軸にそれに対応する残差をとって散布図を画くと図９の如くなる。予測値の小さいところで，残差が多少－側へ傾よっている。

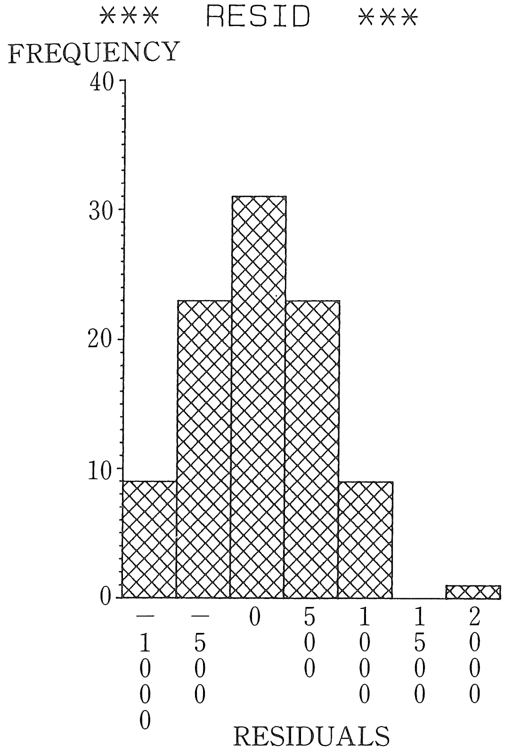

図10　残差の度数分布

５・３　残差の度数分布

残差について度数分布を作ると，図10の如くなる。分散分析の表６より

$Ve=364296.15$

$\sqrt{Ve}=603.57$

が得られ，残差の３シグマ限界を求めると

$\pm3\sqrt{Ve}=\pm1810.71$

で，これを越えるものは異常値とみなされ，No.21の＋2154.6がこれに当てはまる。

５・４　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の値は約２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1.247となり，１％の統計限界の値が約1.445なので，これよりも小さくなっており，正の相関が認められる。

6．おわりに

企業が経営計画を立てる場合や，研究開発を実施するときには，予測が重要なものとなっている。百貨店の売上予測をする場合に，説明変動として消費支出，売場面積，消費者物価指数，時間を用い，昭和54年より61年までのデータをもとに昭和62年の売上を予測した。その結果予測誤差は，最小は７月の－0.050％から最大は４月の6.52％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考えられる。これらについては，さらに検討を進めるつもりである。

参考文献

1）本多正久：BASICによる予測入門，共立出版，東京，1986
2）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
3）愛知県：愛知県統計年鑑，愛知県，1979～1987
4）SAS USER'S GUIDE STATISTICS ver. 5, SAS Institute Inc.
5）SAS USER'S GUIDE GRAPH ver. 5, SAS Institute Inc.
6）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２年２月20日）