大型小売店における売上予測（第３報）

橋本郁郎

Sales Forecast in a Large Size Retail Store (the 3rd Report)

Ikuro HASHIMOTO

A sales forecasting is very important for making a management plan and proceeding development and product planning. There are various forecasts such as technical forecast, economic forecast, sales forecast and product forecast. Succeeding to the previous report, the present study discusess further sales forecast of clothing in the department stores in Aichi Prefecture by means of a multiple regression analysis. In the previous report, such dependent variables were used as consumer expense, sales floor area, time and dummy variables. We know well that department's sales are influenced by temperature. Then the present study added a temperature variable to the dependent variables. As the result of analysis, the autocorrelation of the residual are improved very much in which the forecast error rates are 6.86%max. and −1.57%min.

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・製品計画等を実施するときに，予測が大変有効な手段となっている。第１報[1]では大型小売店のうち愛知県内の百貨店の衣料品総売上高の予測を行ない，第２報[2]では予測手順の中の重回帰分析にダミー変数を導入した。本論ではさらに気温の高低が売上高に及ぼす影響の検討を試みた。

２．研究方法

昭和54年より61年までの月次データ[3]により，昭和62年度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。第２報で述べた予測手順のうちで本研究において異る点は，不規則変動を除くための重回帰分析において，気温の高低の影響を与える説明変数を追加導入したことである。

３．回帰モデル式の構築

気温の高低の売上高に及ぼす影響を見るために，第２報における回帰式に気温差の説明変数を追加して解析することにした。以下の統計解析はＳＡＳ[4)5)]（Statistical Analysis System）によって行なった。

３・１　気温差の導入

月毎の平年気温と平均気温の差が売上高に及ぼす影響を調べるため，(1)式の如き気温差の関数を考える。

$$\left.\begin{aligned} T &= Q \times B \\ B &= TM - TN \end{aligned}\right\} \quad\cdots\cdots(1)$$

Ｔ　：気温差の関数
Ｂ　：気温差　℃
ＴＭ：月平均気温　℃
ＴＮ：月平年気温　℃
Ｑ　：各月毎のウエイトで，夏期と冬期に分け表１の如くＡ，Ｂ，Ｃの３種類を考える

表１　夏期と冬期のウエイト

		夏			期			冬			期		
月		3	4	5	6	7	8	9	10	11	12	1	2
Q	A	1	2	3	4	5	6	1	2	3	4	5	6
	B	1	2	3	4	4	4	1	2	3	4	4	4
	C	1	2	3	4	3	2	1	2	3	4	3	2

気温差の関数を夏期の冬期に分けて考えるため，ダミー変数を導入して(2)式の如き関数を導入する。

$$\left.\begin{aligned} TES &= T \times XS \\ TEW &= T \times XW \end{aligned}\right\} \quad\cdots\cdots(2)$$

ＴＥＳ：夏期気温差の関数
ＴＥＷ：冬期気温差の関数
ＸＳ　：夏期ダミー変数
　３月～８月：ＸＳ＝1，９月～２月：ＸＷ＝0

XW ：冬期ダミー変数

3月～8月：XW＝0，9月～2月：XW＝1

気温差の月次変化を図1に示す。また気温差に対する夏，冬それぞれの気温差の関数TES，TEWを各ウエイトA，B，Cにつき図2～図7に示す。

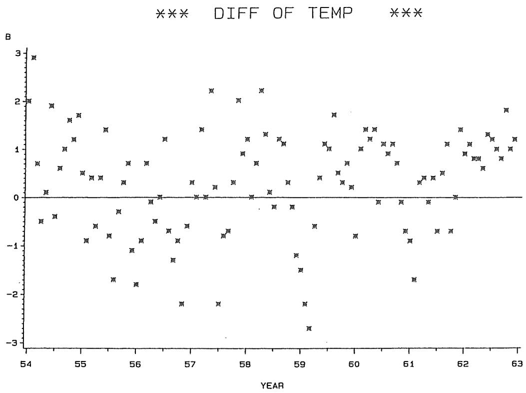

図1　気温差（平均気温－平年気温）

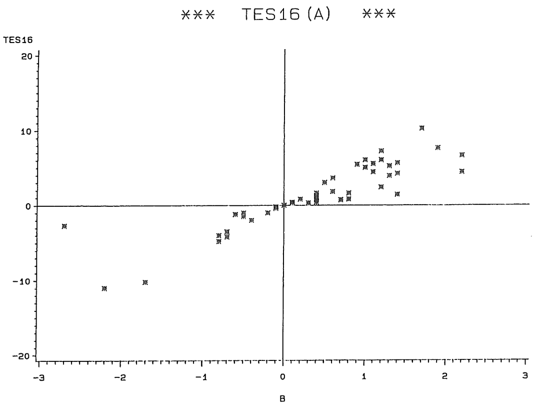

図2　夏期気温差の関数（ウエイトA）

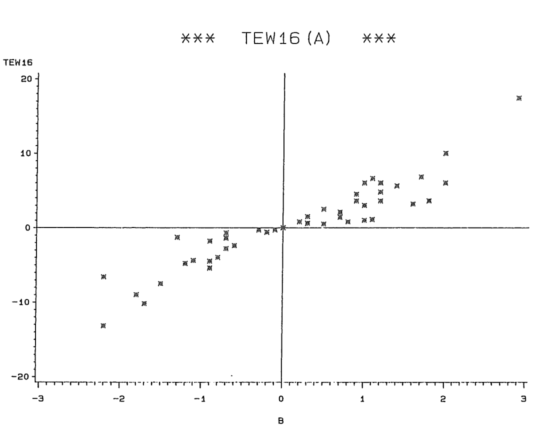

図3　冬期気温差の関数（ウエイトA）

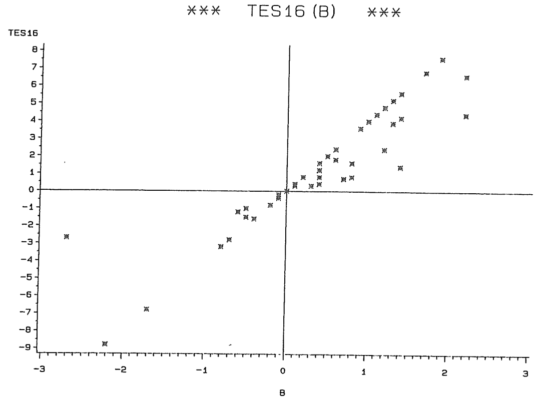

図4　夏期気温差の関数（ウエイトB）

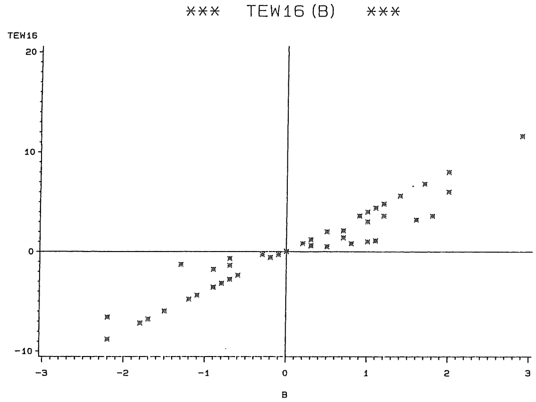

図5　冬期気温差の関数（ウエイトB）

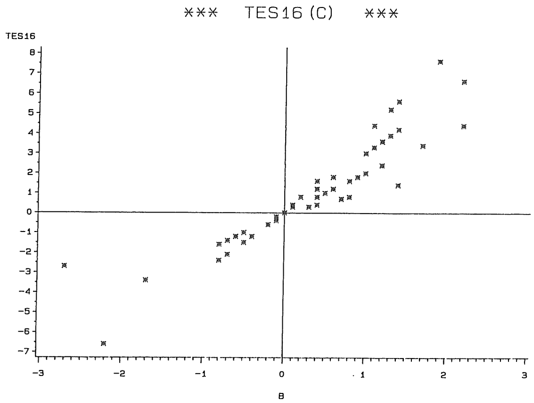

図6　夏期気温差の関数（ウエイトC）

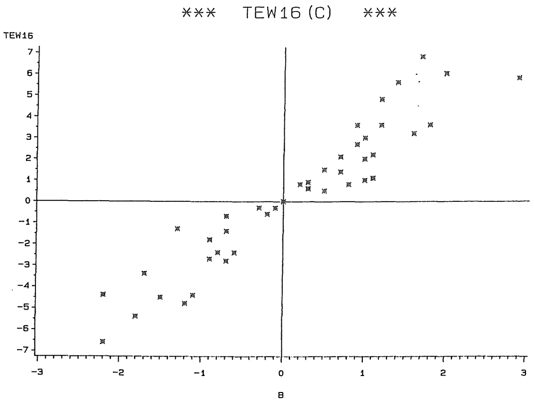

図７　冬期気温差の関数（ウエイトＣ）

３・２　回帰式の構築

第２報のダミー変数を導入した回帰式に気温差の関数を追加し(3)式の如き回帰式により検討を進めた。

$$Y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 \quad \cdots\cdots(3)$$

Ｙ：衣料品の総売上（季調済み）　ＵＲＩＡＧＥ２

x_1：消費支出（季調済み）　ＳＨＩＳＹＵ２

x_2：売場面積　ＭＥＮＳＥＫＩ

x_3：時間　ＴＩＭＥ

x_4：ダミー変数　ＤＵＭ５

x_5：ＤＵＭ５×ＭＥＮＳＥＫＩ

x_6：Ｔ×ＸＳ　ＴＥＳＩ６

x_7：Ｔ×ＸＷ　ＴＥＷＩ６

x_6, x_7については気温差の関数ＴのウエイトＱが３種数あるので，それぞれＴ(Ａ)，Ｔ(Ｂ)，Ｔ(Ｃ)の３組について重回帰分析を行なった。その結果の主なものを表２に示す。この３組につき予測誤差を計算することにする。

表２　重回帰分析結果（気温差の変数）

	VARIABLE	PROB>｜T｜	R-SQUARE	DURBINWATSON
T(A)	TES16	0.6506	0.9263	1.724
	TEW16	0.0048		
T(B)	TES16	0.5651	0.9271	1.715
	TEW16	0.0031		
T(C)	TES16	0.4651	0.9273	1.692
	TEW16	0.0031		

４．予測結果

昭和62年度の売上を予測するには，本来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見る為に実績値を用いることにした。ＴＣ系列の予測値に季節指数[1]をか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。

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

（予測値＝ＴＣ系列の補外値×季節調整指数）

これにより算出した予測値の主なものを表３に示す。Ｔ（Ａ），Ｔ（Ｂ），Ｔ（Ｃ）の間に大差はないが，最大値の一番小さいＴ（Ｂ）をとりあげて回帰モデルの適合性を検討することにする。Ｔ（Ｂ）の場合の予測誤差を表４に示す。

表３　誤差の最大・最小値及び範囲　（％）

	最大値	最小値	範　囲
T(A)	6.96	−1.88	8.84
T(B)	6.86	−1.57	8.43
T(C)	6.97	−1.31	8.28

表４　T(B)の場合の予測誤差　（％）

1　月	2　月	3　月	4　月
5.25	5.44	1.81	6.23
5　月	6　月	7　月	8　月
6.86	2.49	−0.08	2.07
9　月	10　月	11　月	12　月
−1.31	4.46	6.77	−1.57

５．回帰式の検討

(3)　式の説明変数x_6，x_7は気温差の関数のウエイトＱが３種類あるが，このうちＴ（Ｂ）の場合が予測誤差が最小となったので，これを最適モデルと考えて検討を進める。

5・1　重回帰分析結果

最適モデルの重回帰分析結果を表5に示す。またこのＴＣ系列の予測値を図8に示す。

表5　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	7	311483316	44497616.51	159.846	0.0001
ERROR	88	24497197.39	278377.24		
C TOTAL	95	335980513			

ROOT MSE	527.6147	R-SQUARE	0.9271
DEP MEAN	16687.27	ADJ R-SQ	0.9213
C.V.	3.161779		

PARAMETER ESTIMATES

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|
INTERCEP	1	-18099.69736	6540.73052	-2.767	0.0069
SHISYU2	1	0.007099050	0.003585002	1.980	0.0508
MENSEKI	1	89.05466654	19.05496598	4.674	0.0001
TIME	1	35.76662576	6.81276278	5.250	0.0001
DUM5	1	14825.71967	6447.42839	2.299	0.0238
D5MEN	1	-43.76082403	18.17639254	-2.408	0.0181
TES16	1	13.62230628	23.58904446	0.577	0.5651
TEW16	1	-62.09274179	20.42937881	-3.039	0.0031

DURBIN-WATSON D	1.715
(FOR NUMBER OF OBS.)	96
1ST ORDER AUTOCORRELATION	0.116

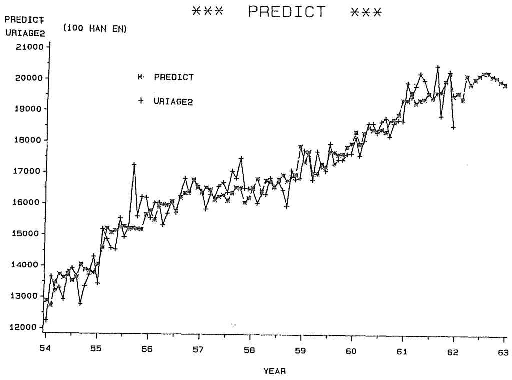

図8　予測値及び実績値（季調済み）

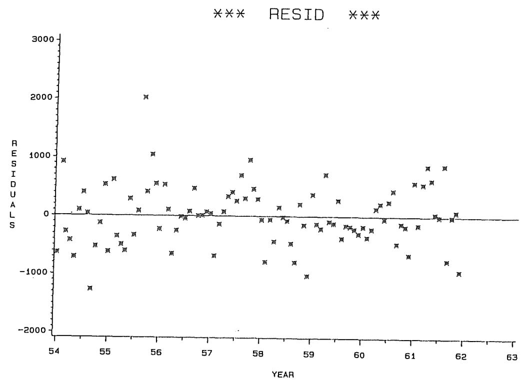

図9　残差の時系列プロット

5・2　残差の検討[7)]

重回帰分析における残差の検討は，回帰式の正当性をみる為に重要である。

5・2・1　残差の時系列プロット

残差の値を時間を横軸にとってプロットすると，図9に示すようになる。残差は零を中心にして上下に変動するが，それにより以下の事項を検討する。

(1)　傾向的な変化

最もマクロ的な見方で，残差プロットに右上り右下りの傾向，周期的な変化，その他曲線的傾向，時間経過による残差の大きさ（絶対値）の変化などはいずれも認められない。

(2)　符号

回帰モデルが正しければ，正符号と負符号は50：50の割合で現われるはずである。本論の場合＋符号n_1が44，－符号n_2が52となっている。これが50：50と見なせるかどうか符号検定表により検定することにする。$n = n_1 + n_2 = 44 + 52 = 96$であるので，$n > 90$

$$F = \frac{n-1}{2} - K\sqrt{n+1} \quad \cdots\cdots\cdots\cdots\cdots(5)$$

　　　K：有意水準１％で　1.2879

　　　　　有意水準５％で　0.9800

の場合は符号検定表の代りに(5)式にて検定する。

本論の場合，有意水準１％のとき

$$F=\frac{96-1}{2}-1.2879\sqrt{96+1}=34.8<$$

$$n_1=44$$

となり，１％の危険率で有意でない。よって残差の符号は50：50と見なせる。

(3)　連の数による検定

＋符号，－符号が50：50と見なせたとしても，その現れ方はランダムでなければいけない。。そのために連の数による検定が必要になる。＋の連，－の連を合計した連の数を n_R とすると，＋，－がランダムに出る場合の 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。

$$\left.\begin{aligned}\mu&=\frac{2n_1n_2}{n_1+n_2}+1\\ \sigma^2&=\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}\end{aligned}\right\}\cdots\cdots\cdots\cdots(6)$$

したがって(7)式の如く u は標準正規分布に従う。

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

これに $n_1=44$，$n_2=52$，$n_R=43$ を代入すると $\mu=48.67$，$\sigma^2=23.415$，$u=1.0678$ となり u は標準正規分布の両側１％点である2.576より小さくなるので有意でない。即ち連の数から見てこの残差系列は，ランダムと見なせる。

(4)　連の長さによる検定

＋，－の符号の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。本論の場合＋側で９個，－側で８個が連続して現われている。７個以上も同じ符号がならぶ場合はランダムとは見ずに，なにかそこに原因があると見た方がよい。

５・２・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10の如くなる。この図には特に問題になるような点は含まれていないと考えられる。

５・２・３　残差の度数分布

残差の度数分布を図11に示す。分散分析の表５より

$$Ve=278377.24$$

$$\sqrt{Ve}=527.61467$$

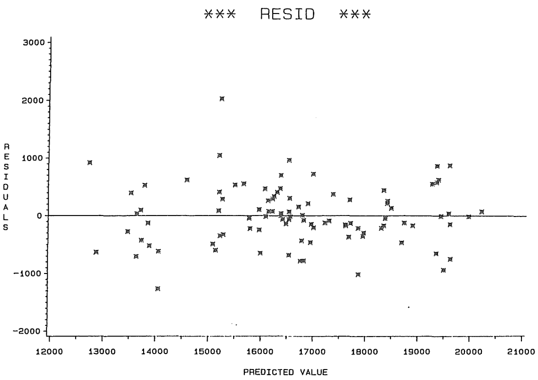

図10　残差と予測値の散布図

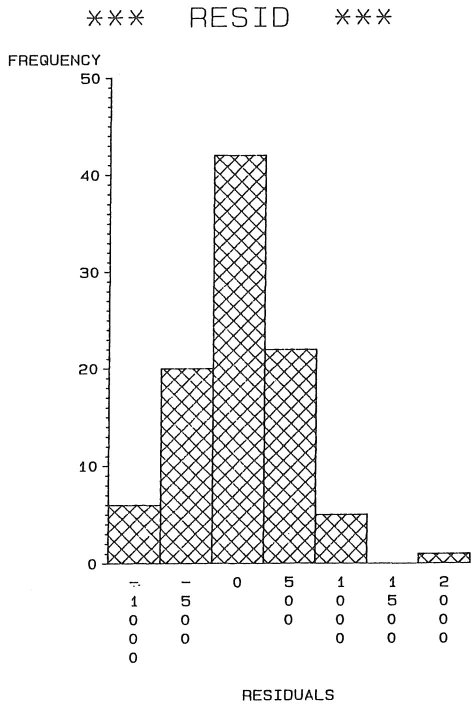

図11　残差の度数分布

となり，残差の３シグマ限界を求めると

$$\pm3\sqrt{Ve}=\pm1582.844$$

で，それを越えるものは№21の2016.61が有り，異状値とみなされる。

５・２・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）を検討する必要がある。誤差の間に相関がなければ，Dの値は２に近づくはずであり，正の相関があればDの値は２より小さく，負の相関があればDの値は２より大きくなる。データの数が100，説明変数が７個

の場合の統計限界は有意水準１％で$D_L=1.40$, $D_U=1.69$と推定される。正の相関を調べるには

$D \leqq D_L$　　正の自己相関あり

$D > D_U$　　正の自己相関なし

$D_L < D < D_U$　　判定出来ない

と検定すればよい。

本論の場合に近似的に適用すると，$D=1.715$なので

$D\ (1.715) > D_U\ (1.69)$

となり，有意な自己相関は認められない。

６．考察

重回帰分析において，第１報のダミー変数なし，第２報のダミー変数を導入した場合と本報のさらに気温差の影響を考慮に入れた場合の比較を表６に示す。これによれば回帰分析全体の信頼性を表す寄与率は90.13％，91.91％，92.71％と向上し，また残差の独立性も大幅に改善され回帰式はより信頼性が高いものになっている。また予測誤差の比較を表７に示す。

表６　寄与率と残差の比較

		第　1　報	第　2　報	第　3　報
寄与率（R^2）		0.9013	0.9191	0.9271
残差	DURBIN-WATSON (D)	1.247	1.622	1.715
	傾向的変化	無　し	無　し	無　し
	符　　号	＋48，−48	＋48，−48	＋44，−52
	連の数による検定	ランダムとは見なせず	ランダムと認められる	ランダムと認められる
	連の長さによる検定	ランダムとは見なせず	ランダムとは見なせず	ランダムとは見なせず

表７　予測誤差の比較　　　（％）

	第１報	第２報	第３報
最大値	6.70	7.13	6.86
最小値	−2.93	−2.82	−1.57
範　囲	9.63	9.95	8.43

７．おわりに

企業が経営計画を立てる場合や，研究開発・製品計画等を実施するときには，予測が重要なものとなっている。百貨店の衣料品の売上予測をする場合は，その年の気温が平年に比して高いか低いかが影響するものと考えられる。そこで気温の影響を気温差の関数として重回帰分析の説明変数に取入れた結果，回帰式の信頼性を向上することが出来た。予測誤差については，最大値6.86％，最小値−1.57％，範囲8.43％となり最大値については第１報よりわずかに高くなったが，最小値，範囲については改善が見られた。

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, PartB, 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）愛知県：愛知県統計年鑑，愛知県，1979～1987

4）SAS USER'S GUID STATISTICS Ver.5, SAS Institute Inc.

5）SAS USER'S GUID GRAPH Ver.5 SAS Institute Inc.

6）名古屋地方気象台：愛知県気象月報，1979～1987

7）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４年３月20日）