<研究ノート>過去原価資料を用いたコスト・ビヘイビアの検討

研究ノート

過去原価資料を用いたコスト・ビヘイビアの分析

加 登 豊

I 序

コスト・コントロールおよびそれらを含む意思決定の種々の局面でコスト・ビヘイビアが明確に把握されていなければ，当該活動の有効性に疑問が生じることになる。コントロール面のみをとりあげても，①予算の作成，②標準原価の設定，③業績報告書の作成，④原価差異の評価など数多くの管理活動は，原価推定値がどれほど信頼できるかによって，その成果が左右されるのである。

それにもかかわらず，コスト・ビヘイビアをあきらかにする方法についての考察は，いまだ十分なものであるとはいえない。そこで本稿では，歴史的原価資料をもとにした原価の推定方法について，コーコランの記述にそって考察する。(1) 本稿で行う説明の大部分は，計量経済学などの分野で広汎に用いられている回帰分析についてのものである。しかし，会計ではそれほど一般的な技法であるとはいえず，会計的解釈を加えねばならないところもあるので，やや冗慢にはなるが，逐一説明を行うことをここでことわっておく。なお，管理技法とコスト・ビヘイビアを明らかにするための技法の間には相互交

＊ 本稿作成にあたって，統計学に関する記述を本学計量経済学講座今川正教授に御検討していただき，有益なコメントを拝聴した。紙面をかりて謝辞を記しておく。

(1) Corcoran [3] pp. 47—87.

コスト・ビヘイビアの把握には一般に，歴史的原価資料を使用する方法，IE 法，および，勘定科目検討法が組合せて行なわれる。それぞれのアプローチの適用範囲，長所，欠点については，Bierman and Dyckman [2] pp. 56—75, Dopuch *et. al.* [4] pp. 50—73, Horngren [6] pp. 777—790, Moore and Jaedicke [10] p. 255—260, 豊島〔21〕pp. 116—126. 等を参照せよ。

渉があるので，それらを一体として考えることが必要であることを付記しておきたい。それでは，最高最低法から議論を出発することにしよう。

Ⅱ　最高最低法

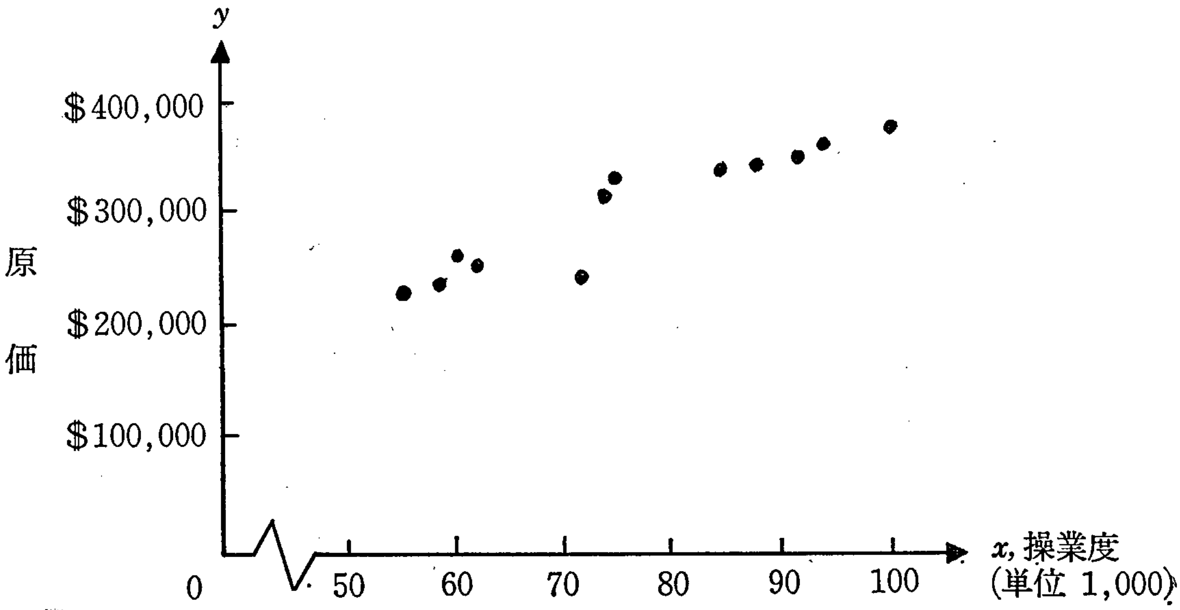
コスト・ビヘイビアを過去の原価資料ならびに操業度に関する資料を基礎として把握しようとする場合，伝統的な方法の１つとして最高最低法（high-low method）がある。これは原価額を，操業度（独立変数あるいは説明変数）変動にあとづけて，どのように変化するかをつかもうとする点で，後述する回帰分析と同一基盤に立つものである。ただ，最高最低法では，周知のように，利用可能な原価ならびに操業度をあらわすデータのうちで，最高操業時と最低操業時のデータ，つまり，２つの観測値をもとにして原価直線をもとめるのである。最高最低法による原価直線の決定方法は，次のようである。なお，本稿で使用するデータは表１にしめされており，図１にそれらのデータがプロットされている。(2)

表　1

	原価（単位1,000ドル）	操業度（単位1,000）			←── 偏差の自乗和	──→	
年	y	x	x^2	xy	$(\bar{y}-y)^2$	$(\bar{y}-\hat{y})^2$	$(\hat{y}-y)^2$
1963	262	60	3, 600	15, 720	1, 806. 25	2, 784. 70	105. 48
1964	230	55	3, 025	12, 650	5, 550. 25	4, 761. 99	30. 17
1965	247	72	5, 184	17, 784	3, 306. 25	190. 48	1, 909. 56
1966	258	62	3, 844	15, 996	2, 162. 25	2, 141. 41	0. 05
1967	240	58	3, 364	13, 920	4, 160. 25	3, 512. 34	27. 41
1968	330	75	5, 625	24, 750	650. 25	16. 48	873. 75
1969	314	74	5, 476	23, 236	90. 25	53. 39	282. 46
1970	340	85	7, 225	28, 900	1, 260. 25	807. 40	50. 20
1971	348	88	7, 744	30, 624	1, 892. 25	1, 455. 95	28. 55
1972	360	94	8, 836	33, 840	3, 080. 25	3, 322. 52	4. 59
1973	350	92	8, 464	32, 200	2, 070. 25	2, 615. 97	31. 88
1974	375	100	10, 000	37, 500	4, 970. 25	5, 948. 37	43. 90
合計（Σ）	$ 3, 654	915	72, 387	287, 120	30, 999. 00	27, 611. 00	3, 388. 00

（2） Corcoran [3] p. 48 and p. 49.

図　1



	年　度	y（原価額）	x（操業度）
最高操業時	1974	$ 375,000	100,000
最低操業時	1964	230,000	55,000
Δ(差異)		$ 145,000	45,000

原価直線の傾き$=\Delta y/\Delta x=\$145,000/45,000=\$3.22=b$

原価直線の方程式　　$y=a+bx=a+\$3.22x$

$x=100,000$ のとき

$$y=\$375,000=a+\$3.22\times 100,000$$

$$\therefore\quad a=\$53,000$$

よって，もとめる原価直線の方程式は

$$y=\$53,000+\$3.22x$$

（固定費）（変動費）

このような，最高最低法による過去原価資料への原価直線のあてはめは，(1)算定処理時間が短かく，(2)理解が容易であるという長所をもつため，実務ではしばしば利用されるコスト・ビヘイビアの把握方法である(3)。しかし，最高最低法によって原価直線を確定しようとする場合には，利用可能な数多くの観測値のうちから，たった2つのデータしか使用しないため，最高操業時および最低操業時のデータが他のすべてのデータを代表するものでない場合には問題がある。データをプロットしたスキャッター・グラフに最高最低法によってもとめた原価直線をえがいてみれば，上記のような状況が存在するか

(3)　*Ibid.*, p. 48.

どうかが直ちに明らかになる。場合によっては，最高最低法でもとめた原価直線をもとにすると，誤った意思決定を下す危険もある。したがって，やむをえず原価直線の決定に２つの観測値を使用するときには，最高最低点を利用することを放棄し，より代表的な他の２点を利用することが望ましい。(4)

さらに，上記の問題にも関連するのであるが，最高最低法やそれに準じる方法によって獲得した原価方程式が，どの程度観測値に適合するものであるかを判断する客観的基準は現在のところ存在しないことに注意すべきである。過去の歴史的原価データを使用して原価方程式を獲得しようとするアプローチでは，最高最低法のみならず後述する回帰分析においても，われわれは暗黙のうちに，過去の原価実績と同様の原価の動きが将来にも考えられることを仮定しているのである。したがって，このようなアプローチでは，まず獲得された原価方程式がその算定の基礎となった観測値に適合するものであるという確証がえられる必要がある。このことが確認されてはじめて，原価方程式を将来にむけて使用することの可能性がひらけてくるのである。

最高最低法のように，簡便なコスト・ビヘイビアの把握方法を使用しようとする場合にでも，次にしめすような困難がある。第１に，操業度をしめす変数（操業度測定単位，これによって原価額を説明しようとする。つまり，原価方程式にふくまれることになる独立変数）に何を選択するかが基本的でかつもっとも重要な問題であるにもかかわらず，実際には，かなり困難な作業なのである。この点については，ここではその指摘だけにとどめ，後に説明を加えることにする。第２に，原価直線の傾き（変動費率）は，選択された観測値いかんによって，相当の変動をまねく結果，固定費額にも重大な影響をあ

（4）　同様の指摘がドパッチらによってもなされている。「この２つの極端な値（最高点・最低点をいう……筆者注）は，ただ全体的な差しか反映しないので，結果としてえられる原価方程式は，いかなる異常状態の存在も考慮していない。最高点および最低点が全体の趨勢をあらわすものでない場合，算定された原価関数は，実際には，正常ではない原価—操業度関係をしめすものとなる。」

Dopuch *et. al.* [4] p. 61.

なお，ホーングレンが注記しているデムスキーによる next high next low method, next-next-high next-next-low method を用いる場合にも，それらの点が観測値を代表する２点であると結論できる根拠は存在しない。ただ，最高・最低点を分析からはずしても，推定量が統計的な種々の特性を失わない状況についての考察が統計学ではなされていることを付記しておこう。Horngren [6] p. 788.

たえることになる。(5) 獲得した原価方程式を，いかなる目的に使用するかによって原価方程式にもとめられる信頼性の程度は異なるのは当然としても，このような不明瞭なコスト・ビヘイビアの把握方法にもとづく原価方程式を，種々の意思決定問題に適用するということ自体，相当な問題を内包しているものといえる。したがって，最高最低法は，いかに簡便であるとはいえ，その使用はつつしむべきである。

Ⅲ　回　帰　分　析

3-1　概略説明

過去の原価資料をもとにして，まず，原価がどのように動いてきたかを分析することが歴史的原価資料を用いてコスト・ビヘイビアを明らかにするためには必要である。さらに，将来意思決定を行うさいに，参考となる未来原価に関する資料が獲得できないとか，信頼できる資料がえられないような場合，一定の制約条件つきで，過去原価の動向をしめす情報が重要となる。このように，過去原価に関する資料は，言葉をかえていえば，経営活動のコントロール面のみならずプランニングの局面にも使用されるのである。

特に，プランニングに過去原価資料をもとにした原価方程式を使用しなければならないときには，できる限り，当該原価方程式が信頼できるものでなければならない。原価方程式が不正確であれば，意思決定にともなう幾多の判断が正しくても，良好な結果はえられないだろう。このことに関して，後述する回帰分析の有用性が指摘できる。ただし，原価方程式の確定に回帰分析という統計的方法を適用する場合，次の警句を決して忘れてはならない。

「独立変数と従属変数の間に一定の因果関係が存在する場合，それは統計的分析に組込まれるものであり，分析によって明らかにされるという性質のものではない。つまり，回帰分析がなしうるのは，どのように変数どうしが関連するかを叙述することである。(6)」

したがって，原価方程式の確定に統計的分析の適用を考える状況を想定すれば，分析

(5)　Corcoran [3] p. 50.

(6)　*Ibid.*, p. 51.

(7)　操業度測定単位に何を選択するかは，独立変数が従属変数とどの程度関連するかという事項とは，一応，別個の問題であると理解するべきである。このような認識↗

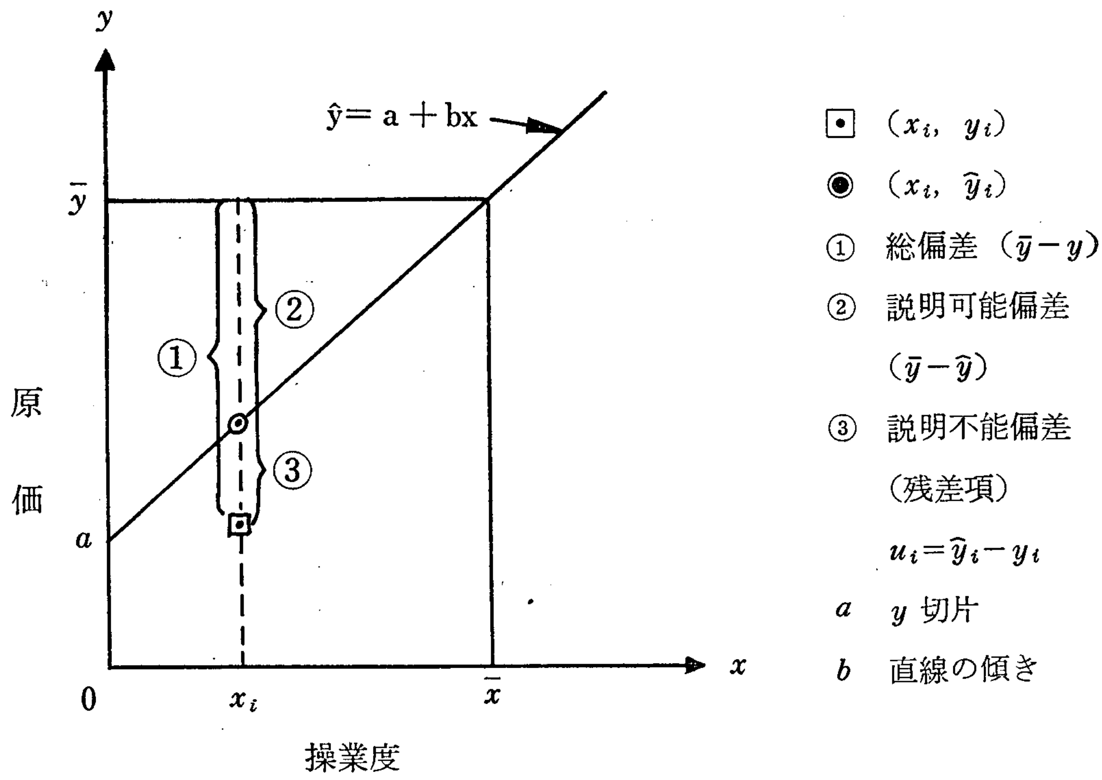
者は分析を行うにあたって，独立変数と従属変数の間に，後に分析を通じて使用されるであろう原価方程式が種々の経営管理目的に対し合目的的であるように，一定の因果関係を設定しておかねばならない。[7]

次節以降で回帰分析手法の概略を説明する。回帰分析の結果を誤まりなく，かつ有効に利用するためには，先にものべたが分析実施の準備段階で，過去の操業度に関するデータから，組織が従来，どの程度の範囲で操業してきたかを認識しておくことが重要である。さらには，回帰方程式からもとめられる推定値，観測値，過去原価の平均値を比較すると同時に，それらの過去の趨勢を把握しておくことも重要である。

3-2　最小自乗法によるパラメータの推定

まず，コーコランは図2を用いて推定原価直線を確定する手順について記述している。

図　2



が分析者にないと，正しい分析手順をふんで獲得された結果の解釈を誤まる危険がある。2つの変数が強く関連する（strongly associated）ことを統計学の用語法によれば，高い正の相関をしめす（exhibit a high, positive correlation）という。再度のべるが，変数の関連の度合を決定することは相関分析（correlation analysis）の範疇に属すること（*Ibid.*, p. 51.）であり，変数間の因果関係とは無関係なのである。

図2で y_i は期間 i の実際原価観測値，$\bar{y}$ は平均原価，$\hat{y}_i$ は期間 i の原価の回帰推定値である。回帰直線をもとめる過程では，単に推定原価方程式 $\hat{y}=a+bx$ に含まれるパラメータの推定値，a および b（a は定数項 constant term，b は回帰係数 regression coefficient とよばれる）を確定するだけであるから，回帰分析に関する統計的諸仮定については本節では言及しない。推定原価方程式に特定の x の値を代入すれば，その x に対する y の真の平均値の予測値がえられることになる。観測値（y_i）が獲得される過程では，「どのような観測においても，ある程度の測定誤差はどうしても避けられないものである。(8)」そこで，実際測定値（観測値）は次のようにあらわすことができる。

$$y_i=a+bx_i+u_i \qquad i=1,\ 2,\ \cdots\cdots,\ n \tag{1}$$

ここで u_i は撹乱項である。(9)

パラメータの推定値 a，b を獲得する方法の1つに最小自乗法がある。これは撹乱項の自乗の和，

$$u_1{}^2+u_2{}^2+u_3{}^2+\cdots\cdots+u_n{}^2=\sum_{i=1}^{n}u_i{}^2 \tag{2}$$

を最小にするように，パラメータの推定値 a，b を決める方法である。(10)観測値の組の数を n，撹乱項の自乗和を S とすれば，

$$S=\sum_{i=1}^{n}u_i{}^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\sum_{i=1}^{n}(y_i-a-bx_i)^2 \tag{3}$$

S を最小にするパラメータの推定値 a，b の値は，次の条件をみたすことが必要である。

$$\begin{cases}\dfrac{\partial S}{\partial a}=-2\displaystyle\sum_{i=1}^{n}(y_i-a-bx_i)=0\\[2ex]\dfrac{\partial S}{\partial b}=-2\displaystyle\sum_{i=1}^{n}x_i(y_i-a-bx_i)=0\end{cases} \tag{4}$$

（8） 岩田〔14〕p. 188.

（9） 撹乱項 u_i は，

$$u_i=y_i-a-bx_i \qquad i=1,\ 2,\ \cdots\cdots,\ n$$

とすると，未知のパラメータの推定値 a，b の関数であると考えられる。確率的撹乱項（random disturbance）は，統計的分析をさらに展開する上でいくつかの仮定のもとにある確率変数である。撹乱項にかかわる諸仮定については後に詳説する。

（10） 岩田〔14〕pp. 189—190.

(4)式を書きあらためると，

$$\begin{cases} na+b\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i \\ a\sum_{i=1}^{n}x_i+b\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_iy_i \end{cases} \tag{5}$$

(5)式は，回帰直線の正規方程式（normal equation of the regression line）とよばれるものである。

正規方程式をマトリクス表示すると次のようになる。(11) 以下，添字は表記法を簡略化するために省略する。

$$\begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix} \tag{6}$$

$\boldsymbol{X}$を独立変数のマトリクス（$n\times2$），$\boldsymbol{Y}$を観測値のベクトル（$n\times1$），$\boldsymbol{B}$をパラメータの推定値のベクトル（2×1）とする。つまり，

$$\boldsymbol{X}=\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \qquad \boldsymbol{Y}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad \boldsymbol{B}=\begin{bmatrix} a \\ b \end{bmatrix}$$

$\boldsymbol{X}$の転置行列を$\boldsymbol{X}^T$とすれば，(6)式は次のようにあらわされる。

$$(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{B}=(\boldsymbol{X}^T\boldsymbol{Y}) \tag{7}$$

(7)式を解けば，

$$\underbrace{(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{X})}_{\boldsymbol{I}}\boldsymbol{B}=(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{Y})$$

$$\boldsymbol{B}=(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{Y}) \tag{8}$$

となる。

正規方程式を解く手順を表1の数値を並記しながら見ていくことにしよう。

---

(11)　マトリクス表示の利点は，回帰分析に複数の変数が導入されても（節3－9の重回帰分析の項をみよ），正規方程式の表示法が変わらないという計算技術上の利点である。Corcoran [3] p. 54.

同様の主張は次の文献にもみられる。

中村訳〔17〕p. 47.

$$
\begin{array}{ccccc}
\boldsymbol{X}^T & & \boldsymbol{X} & & \boldsymbol{X}^T\boldsymbol{X} \\
\begin{bmatrix} 1 & 1 \cdots\cdots 1 \\ x_1 & x_2 \cdots\cdots x_n \end{bmatrix} & & \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} & = & \begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{bmatrix}
\end{array}
$$

（8）式の左側のマトリクス

$$
\begin{array}{ccccc}
\boldsymbol{X}^T & & \boldsymbol{Y} & & \boldsymbol{X}^T\boldsymbol{Y} \\
\begin{bmatrix} 1 & 1 \cdots\cdots 1 \\ x_1 & x_2 \cdots\cdots x_n \end{bmatrix} & & \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} & = & \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix}
\end{array}
$$

（8）式の右側のベクトル

$$
\begin{array}{ccccc}
\boldsymbol{B} = & & (\boldsymbol{X}^T\boldsymbol{X})^{-1} & & \boldsymbol{X}^T\boldsymbol{Y} \\
\begin{bmatrix} a \\ b \end{bmatrix} & = & \dfrac{\begin{bmatrix} \Sigma x^2 & -\Sigma x \\ -\Sigma x & n \end{bmatrix}}{n\Sigma x^2 - (\Sigma x)^2} & & \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix}
\end{array}
$$

$$
= \frac{\begin{bmatrix} \Sigma x^2 \Sigma y - \Sigma x \Sigma xy \\ n\Sigma xy - \Sigma xy \end{bmatrix}}{n\Sigma x^2 - (\Sigma x)^2}
$$

表1に示されている原価と操業度についてのデータを代入すると

$$
\begin{array}{ccccc}
\boldsymbol{X}^T & & \boldsymbol{X} & & \boldsymbol{X}^T\boldsymbol{X} \\
\begin{bmatrix} 1 & 1 \cdots\cdots 1 \\ 60 & 55 \cdots\cdots 100 \end{bmatrix} & & \begin{bmatrix} 1 & 60 \\ 1 & 55 \\ \vdots & \vdots \\ 1 & 100 \end{bmatrix} & = & \begin{bmatrix} 12 & 915 \\ 915 & 72,387 \end{bmatrix}
\end{array}
$$

$$
\begin{array}{ccccc}
\boldsymbol{X}^T & & \boldsymbol{Y} & & \boldsymbol{X}^T\boldsymbol{Y} \\
\begin{bmatrix} 1 & 1 \cdots\cdots 1 \\ 60 & 55 \cdots\cdots 100 \end{bmatrix} & & \begin{bmatrix} 262 \\ 230 \\ \vdots \\ 375 \end{bmatrix} & = & \begin{bmatrix} 3,654 \\ 287,120 \end{bmatrix}
\end{array}
$$

$$\boldsymbol{B} = \quad (\boldsymbol{X}^T\boldsymbol{X})^{-1} \quad \boldsymbol{X}^T\boldsymbol{Y}$$

$$= \frac{\begin{bmatrix} 72,387 & -915 \\ -915 & 12 \end{bmatrix}}{12(72,387)-915^2} \begin{bmatrix} 3,654 \\ 287,120 \end{bmatrix}$$

$$= \frac{\begin{bmatrix} 72,387 & -915 \\ -915 & 12 \end{bmatrix}}{31,419} \begin{bmatrix} 3,654 \\ 287,120 \end{bmatrix}$$

$$= \begin{bmatrix} 2.3039243 & -0.0291225 \\ -0.0291225 & 0.0003819 \end{bmatrix} \begin{bmatrix} 3,654 \\ 287,120 \end{bmatrix}$$

$$= \begin{bmatrix} 56.8859 \\ 3.2472 \end{bmatrix}$$

以上の結果から，推定原価直線の方程式は次のようになる。

$$\hat{y} = \$56,885.9 + \$3.2472x$$

（固定費）　（変動費）

上記の手順でもとめられる a，b の値は，最小自乗法という1つの基準により獲得される推定値であることに注意しなければならない。[12] 最小自乗法によって獲得されるパラメータの推定値 a，b の値は，S の値を最小化する。このことは容易に証明できる。[13] a，b のもとめられた値が点推定値（point estimates）として統計的にみてすぐれていることが最小自乗法を回帰分析に適用する最大の理由である。[14] 推定値が信頼できるものであ

(12)　岩田〔14〕p. 191.

(13)　*Ibid.*, p. 191，脚注2．をみよ。

(14)　点推定の良さをしめす条件とは次のものをいう。

(i)　一致性をみたすこと。

(ii)　不偏性をみたすこと。

(iii)　有効性をみたすこと。

(iv)　充足性をみたすこと。

最小自乗法は上記の条件をみたす最も簡潔な点推定法であると一般には理解されている。藤澤，松行〔12〕pp. 192—194.

最小自乗法が上記の条件をみたすことについては，後節で回帰分析の諸仮定を列挙する際に説明を加えることにする。

ることが知られていれば，それにもとづいて意思決定を行う際に，原価方程式自体に無用の考慮を払うことがなくなる。このような意味で，統計的分析の知識も会計では不用のものであるとはいえない。われわれの目的は，原価方程式を確定することのみにあるのではなく，それを基礎としてマネジメントの種々の局面にインパクトをあたえることにもある。(15) それゆえ，統計的方法のすぐれている点とその適用上の限界の両者について熟知していなければならない。

3-3　推定回帰線のあてはまり具合の測定（決定係数）

回帰方程式は観測値にどれほど適合するものであろうか。回帰方程式のあてはまりのよさは，全変動（total deviation）に対する回帰変動（explainable deviation）の比率で計算される決定係数（coefficient of determination）；r^2 で測定される。(16) 決定係数は

---

(15)　かつて，最小自乗法の適用に対して次のような批判がなされたことがある。「たいていの場合，この直線はスキャッター・グラフ法によって最もよくあてはめられる。最小自乗法は図にある点の特定の集合により正確にあてはまる直線をもたらすけれども，もとの原価数字がおおざっぱな近似値にすぎないので，このように正確にみえるのは虚構であろう。……インタビューしたある会社は両方の方法をテストして比較したが，この2つの方法の正確性に重大な相違はみられなかったと述べた。」（NAA, Research Report No. 16, 17, 18, *The Analysis of Cost-Volume-Profit Relationships*, 1949—50, p. 16, *fn.* 12.）このように分析の基礎資料が不正確である場合には，統計的方法の適用は厳に慎しむべきである。（豊島〔21〕p. 122 脚注24参照）しかしながら，これを理由に統計的方法の適用を放棄すべきではない。分析される資料をできうる限り正確に測定する努力をおしんではならない。正確に測定された資料をもとにすれば，最小自乗法によって資料にもっともよく適合する原価直線が一意的にえられるだけではなく，従来，会計の分野ではみすごされがちであった獲得された結果の信頼性をも測定すること（統計的検定）が可能となるのである。最小自乗法によってえられた原価方程式をマネジメントの種々の局面で適用する場合にはその信頼性に関する知識は不可欠なのである。

(16)　全変動とはすべての観測値についての総偏差（$\bar{y}-y_i$）の自乗和である（図2参照）。同様に回帰変動は説明可能偏差（$\bar{y}-\hat{y}_i$）の自乗和，誤差変動は説明不能偏差（$\hat{y}_i-y_i$）の自乗和である。換言すれば，全変動，回帰変動，誤差変動はそれぞれ，観測値の平均よりの偏差の平方和，推定値の平均からの偏差の平方和，観測値の推定値のまわりの偏差（つまり残差）の平方和，である。なお，　↗

次式から明らかなように，y の変動のうちでどの程度が，回帰方程式によって説明されるかをしめす係数である。

$$r^2=\frac{\text{回帰変動}}{\text{全変動}}=\frac{\Sigma(\bar{y}-\hat{y})^2}{\Sigma(\bar{y}-y)^2} \qquad (9)$$

われわれは，決定係数を判断基準として，よりあてはまり具合のよい回帰方程式をもとめる方法をとるか，現状で満足すべきかを決定する。(9)式はまた，次のようにも書きあらわすことができる[17]。

$$\Sigma(\bar{y}-y)^2=\Sigma(\bar{y}-\hat{y})^2+\Sigma(\hat{y}-y)^2$$

（全変動）（回帰変動）（誤差変動）

が成立つことは注17をみよ。

上式から $0\leq r^2\leq 1$ であることは明らかであろう。

(17)　(9)式から(10)式は次のようにして導かれる。

$\hat{y}$ は x に値があたえられた場合の y の推定値であるから

$$\hat{y}=a+bx \qquad ①$$

(5)式でしめされる正規方程式により

$$\Sigma y=na+b\Sigma=\Sigma(a+bx)=\Sigma\hat{y} \quad \text{（①式より）} \qquad ②$$

$$\Sigma xy=a\Sigma x+b\Sigma x^2 \qquad ③$$

また，$\Sigma y\hat{y}=\Sigma y(a+bx)=a\Sigma y+b\Sigma xy$　④

$$\begin{aligned}&\Sigma(\bar{y}-\hat{y})^2+\Sigma(\hat{y}-y)^2\\&=\Sigma(\bar{y}^2-2\bar{y}\hat{y}+\hat{y}^2)+\Sigma(\hat{y}^2-2\hat{y}y+y^2)\\&=\Sigma\bar{y}^2-2\bar{y}\Sigma\hat{y}+2\Sigma\hat{y}^2-2\Sigma y\hat{y}+\Sigma y^2\end{aligned}$$

ここで，④式をさらに展開すると，

$$\begin{aligned}\Sigma y\hat{y}&=a(na+b\Sigma x)+b(a\Sigma x+b\Sigma x^2) \quad \text{（②，③式より）}\\&=na^2+2ab\Sigma x+b^2\Sigma x^2\\&=\Sigma(a^2+2abx+b^2x^2)\\&=\Sigma\hat{y}^2 \qquad ⑤\end{aligned}$$

したがって，

$$\begin{aligned}&\Sigma(\bar{y}-\hat{y})^2+\Sigma(\hat{y}-y)^2\\&=\Sigma\bar{y}^2-2\bar{y}\Sigma y+2\Sigma\hat{y}^2-2\Sigma y\hat{y}+\Sigma y^2 \quad \text{（②式より）}\\&=\Sigma\bar{y}^2-2\Sigma\bar{y}y+2\Sigma\hat{y}^2-2\Sigma\hat{y}^2+\Sigma y^2 \quad \text{（⑤式より）}\\&=\Sigma(\bar{y}^2-2\bar{y}y+y^2)\\&=\Sigma(\bar{y}-y)^2\end{aligned}$$

$$r^2=1-\frac{\Sigma u^2}{\Sigma(y-\bar{y})^2} \tag{10}$$

つまり，残差平方和が小さくなると，決定係数の値が1に近づくのである。残差平方和が小さいほど，回帰直線のあてはまり具合がよいと考えて，r^2 を回帰直線のあてはまりの程度をしめす指標として使用するのである。表1の数値で設例の決定係数をもとめると，

$$r^2=\frac{\Sigma(\bar{y}-\hat{y})^2}{\Sigma(\bar{y}-y)^2}=\frac{27,611}{30,999}=0.8907061 \tag{11}$$

となる。この決定係数の値から，回帰方程式はかなりよく観測値にあてはまっているといえる。しかし，このようなあてはまり具合のよさも独立変数と従属変数に共通する増加傾向によることも考えられるので，決定係数が絶対的なあてはまり具合の基準ではないことに注意しなければならない。

後の節で分析をさらに進めるため，ここでは，決数係数はまた，次のようにも書きあらわすことができることをしめしておく。(18)

↙つまり $\Sigma(\bar{y}-\hat{y})^2+\Sigma(\hat{y}-y)^2=\Sigma(\bar{y}-y)^2$ が成立つ。
（回帰変動）（誤差変動）（全変動）

そこで

$$r^2=\frac{\Sigma(\bar{y}-\hat{y})^2}{\Sigma(\bar{y}-y)^2}=\frac{\Sigma(\bar{y}-y)^2-\Sigma(\hat{y}-y)^2}{\Sigma(\bar{y}-y)^2}=1-\frac{\Sigma(\hat{y}-y)^2}{\Sigma(\bar{y}-y)^2}$$

定義より $\hat{y}-y=u$，したがって

$$r^2=1-\frac{\Sigma u^2}{\Sigma(\bar{y}-y)^2}$$

Corcoran [3] pp. 56—57.

(18)　(12)式導出の手順は次のとおり，

$$r^2=\frac{\Sigma(\bar{y}-\hat{y})^2}{\Sigma(\bar{y}-y)^2}=\frac{\Sigma(\bar{y}-\hat{y})^2/n}{\Sigma(\bar{y}-y)^2/n}=\frac{\Sigma(\bar{y}-\hat{y})^2/n}{V(y)}$$

ここで，$\frac{1}{n}\Sigma(\bar{y}-\hat{y})^2=\frac{1}{n}\Sigma(\hat{y}-\bar{y})^2$

$$=\frac{1}{n}(a+bx-\bar{y})^2$$

$$=\frac{1}{n}\Sigma(a+bx)^2-\frac{2}{n}\Sigma(a+bx)\bar{y}+\frac{1}{n}\Sigma\bar{y}^2$$

$$=a^2+2ab\bar{x}+b^2(\Sigma x^2/n)-2a\bar{y}-2b\bar{x}\bar{y}+\bar{y}^2$$

$$=(a-\bar{y})^2+2b\bar{x}(a-\bar{y})^2+b^2(\Sigma x^2/n)$$ ↗

$$r^2=\frac{[\mathrm{Cov}(x,\ y)]^2}{V(x)V(y)} \tag{12}$$

$V(x)$，$V(y)$ はそれぞれ，x と y の分散，$\mathrm{Cov}(x,\ y)$ は x と y の共分散である。

3－4　回帰の標準偏差

前節で説明した決定係数 r^2 の値が大きい（つまり，1に近づく）ということは，ある意味では，回帰分析を用いてわれわれが行った原価分析が正確であることをしめすものである。[19] しかし，回帰直線の方程式の観測値へのあてはまり具合は，さらに，次のような観点からも検討される必要がある。すなわち，観測値が回帰直線のまわりでそのちらばり具合が小さいほど，その回帰直線はデータに適合するものであるということである。このことを測定する統計量は，回帰の標準偏差（standard deviation of regression）である。[20] つまり，回帰の標準偏差は，回帰直線のまわりにみられる実際観測値のちらばりの測度である。とくに回帰方程式を原価予測に使用する状況を考える場合には，回帰の標準偏差の値ができうるかぎり小さいことが望ましい。

回帰の標準偏差は $S_{y\cdot x}$ であらわされる。これは，観測誤差分散の平方根に等しい。

$$=[(a-\bar{y})+b\bar{x}]^2-b^2\bar{x}^2+b^2(\Sigma x^2/n)$$

正規方程式の第一式を使用し，共通項をくくりだして式を整理すると，$\bar{y}=a+b\bar{x}$ であるから，

$$\frac{1}{n}\Sigma(\bar{y}-\hat{y})^2=0^2+b^2(\Sigma x^2/n-\bar{x}^2)$$

$$=b^2V(x)$$

ここで，$b=\dfrac{n\Sigma xy-\Sigma x\Sigma y}{n\Sigma x^2-(\Sigma x)^2}=\dfrac{n^2\mathrm{Cov}(x,\ y)}{n^2V(x)}=\dfrac{\mathrm{Cov}(x,\ y)}{V(x)}$

したがって，$\dfrac{1}{n}\Sigma(\bar{y}-\hat{y})^2=\dfrac{[\mathrm{Cov}(x,\ y)]^2}{V(x)}$

ゆえに，$r^2=\dfrac{[\mathrm{Cov}(x,\ y)]^2}{V(x)V(y)}$

Corcoran [3] pp. 57—58.

(19)　*Ibid.*, p. 58.

(20)　回帰の標準偏差とは一般に，推定値の標準誤差（standard error of the estimate）と呼ばれているものである。コーコランは後者の呼称が当該統計値を誤って使用する原因となっているとし，その使用を戒めている。
Ibid., p. 58.

$$S_{y \cdot x}=\sqrt{\frac{\Sigma(\hat{y}-y)^2}{n}}=\sqrt{\frac{3,388}{12}}=16.80 \tag{13}$$

同様に，従属変数（被説明変数）にのみ注目すれば，標本標準偏差（sample standard deviation）S_y——設例の場合には，原価変動の平方根である——も算出できる。

$$S_y=\sqrt{\frac{\Sigma(\bar{y}-y)^2}{n}}=\sqrt{\frac{30,999}{12}}=50.83 \tag{14}$$

自由度を考慮して(13)式，(14)式でしめされる統計量を調整すればつぎのように回帰の標準偏差推定値がえられる。

$$\hat{S}_{y \cdot x}=\sqrt{\frac{n}{n-2}}S_{y \cdot x}=16.80\times\sqrt{\frac{12}{10}}=18.41 \tag{15}$$

$$\hat{S}_y=\sqrt{\frac{n}{n-1}}S_y=50.83\times\sqrt{\frac{12}{11}}=53.09 \tag{16}$$

調整された標準偏差の値を用いて，われわれはコントロール・チャート（production control chart）[21]を作成する手順とほぼ同一の方法で，推定値のまわりの信頼限界（confidence limits around estimates）を設定することができる。「推定値のまわりの信頼限界によって，われわれは，推定値がわれわれが望ましいと思うだけの正確性（精度，つまり許容性）と信頼性をもつと言明することが可能となるのである。[22]」

観測値 y_i と残差項 u_i が正規分布すると仮定すれば（あるいは，正規分布すると信じるにたる根拠があれば），統計的観点から，一定の確率で，観測値が回帰直線のまわりの一定の範囲内に存在するということができる。観測値数が小さいので，これを t 分布にしたがうものと見なして，次式により，一定の確率でデータが回帰直線のまわりに存在する範囲（特定の x の値について，y が一定の確率を付与された場合にとりうる値の範囲），すなわち，信頼区間帯（confidence band）を算定できる。[23]

---

(21)　コントロール・チャートについては多くの文献があるが，たとえば次のものを参照せよ。

神戸大学会計学研究室編「管理会計ハンドブック」中央経済社，昭和44年，pp. 989—999.，Corcoran [3] pp. 430—433.，Dopuch *et. al.* [4] pp. 484—486.，Boot, J. C. G. and E. B. Cox, *Statistical Analysis for Managerial Decisions*, McGraw-Hill, 1974, pp. 522—534.

(22)　Corcoran [3] p. 59.

(23)　(17)式の導出については，中村訳〔17〕pp. 23—24. をみよ。

$$\hat{y} \pm t_{n-2,\alpha/2} \times \hat{S}_{y \cdot x} \times \sqrt{\frac{1}{n}+\frac{(x-\bar{x})^2}{\Sigma(x-\bar{x})^2}} \tag{17}$$

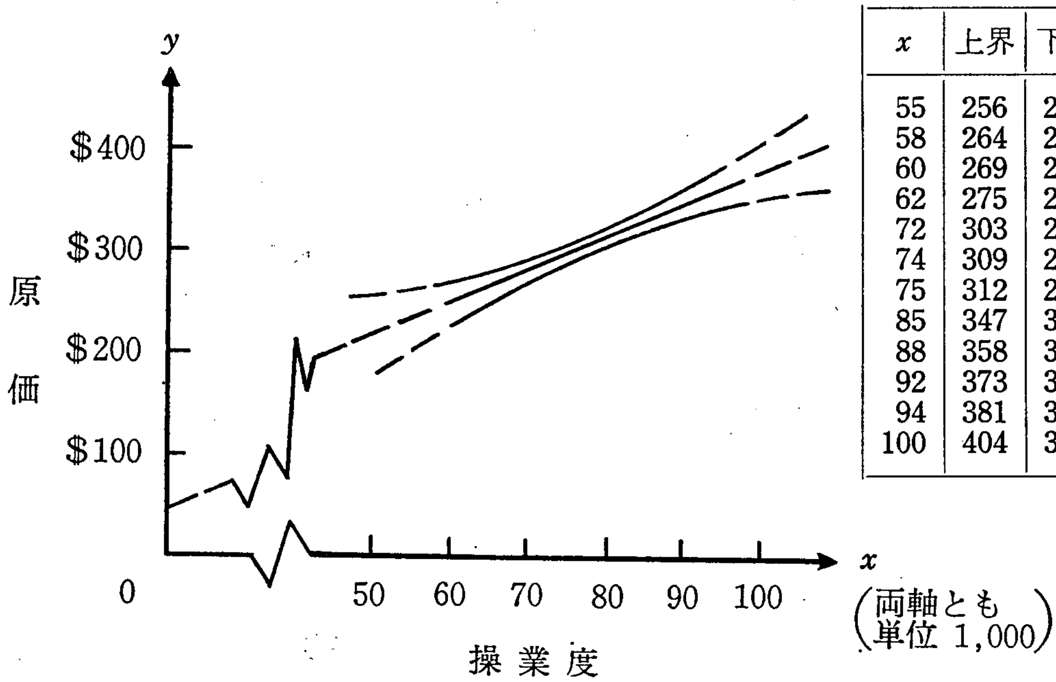
ここで，$t_{n-2,\alpha/2}$ は t 分布表からのよみである。(n……観測値数，α……特定の x の値に対してそれに対応する y の値が信頼区間内にふくまれない確率)

表1のデータで最小自乗法によって算定した回帰直線の95%信頼区間を設定すると，その信頼区間帯は次のようになる。

$$(56,885.90+3.2474x) \pm 2.228 \times 16.41 \times \sqrt{\frac{1}{12}+\frac{(x-76.25)^2}{2,618.25}} \tag{18}$$

(18)式をグラフでしめすと図3のようになる。信頼区間帯が，標本平均 $\bar{x}$ のところでも

図　3



x	上界	下界
55	256	215
58	264	226
60	269	234
62	275	242
72	303	278
74	309	285
75	312	289
85	347	319
88	358	328
92	373	338
94	381	344
100	404	359

っとも狭くなり，標本平均から離れるのにしたがって幅広くなるような放物線であらわされることに注目してほしい。これは，(17)式から明らかなように，独立変数 x の値によって，信頼区間の上界，下界の値が変化するからである。[(24)] このことは，最小自乗法と

(24)　平均のまわりの分散 (variance around an average) は個々の観測値のまわりの分散より小さいので，統計学者は個々の y の予測値の信頼限界を次のように算定する。

$$\hat{y} \pm t_{n-2,\alpha/2}\hat{S}_{y \cdot x}\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{\Sigma(x-\bar{x})^2}}$$

↗

いう基準によってサンプルにあてはまり具合のよい回帰直線が導出され，かつ，推定されたパラメータ値も統計的にすぐれたものであっても，当該原価方程式を用いて原価額の予測を行う場合には，独立変数の値が標本平均から離れていけばいくほど，回帰直線の信頼性が薄れてくることを意味している。なお，ここでは，回帰の標準偏差値を一定としてある。このような状況は，分散斉一性の仮定（assumption[25] of homoscedasticity）とよばれるが，次節でその意味を検討したい。

3-5　回帰分析の諸仮定とその会計的観点からの検討

回帰直線の方程式を確定するさいには，何らの統計学的観点からの仮定も必要ではなかった。しかし，推定値 $\hat{y}$ の信頼限界を算出するためには，観測値 y ならびに残差項 u が正規分布するという仮定をおいた。このように，さらに回帰方程式に対し有用な統計的言明を行うには，種々の仮定をおく必要がある。これらの仮定が維持されている場合にかぎり，回帰分析によって母集団の特性を，そこから抽出されたサンプルを通じて把握することが可能となるのである。そこで，回帰分析の諸仮定を列挙した後に，それぞれについて若干のコメントを付記することにする。[26]

予測値のまわりの分散と平均のまわりの分散のちがいは次にしめすとおり。すなわち，$\hat{y}$ のまわりの信頼限界は，標本 y_t が回帰線を決定する時に用いたのと同じ固定値 x のところで同じ大きさでとられたら，その特定の x に対する y の平均値が95％の確率で真の母集団からの回帰線がこの限界内に存在することをしめすのに対し，$\hat{y}$ のまわりの信頼限界は，個々の観測値 y が対応する所与の x の値について，95％の確率でその信頼区間内にあるだろうということをしめすものである。

Ibid., p. 60, *fn.* 2.

(25)　「分散斉一性」という訳語は今川〔13〕によっている。

(26)　Corcoran ［3］ pp. 60—62.

それではここで注14の補足説明を行うことにしよう。

まず，不偏推定量とは，推定値の平均が，パラメータの平均に等しい推定量である。独立変数と従属変数の関係が次のようにあらわされるとする。

$$y=\alpha+\beta x+u$$

$\bar{y}=\alpha+\beta\bar{x}$ であるから，回帰係数の推定値 b は，

$$b=\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

〔仮定１〕 x（独立変数）と y（従属変数）の関係は，直線によってあらわすことが

$$=\frac{\Sigma(x-\bar{x})\{\beta-(x-\bar{x})+u\}}{\Sigma(x-\bar{x})^2}$$

$$=\frac{\beta\Sigma(x-\bar{x})^2+\Sigma u(x-\bar{x})}{\Sigma(x-\bar{x})^2}$$

$$=\beta+\frac{\Sigma u(x-\bar{x})}{\Sigma(x-\bar{x})^2}$$

$E(b)=\beta$, したがって，b は β の不偏推定量である。同様に $E(a)=\alpha$ となるから a も α の不偏推定量である。

一致性とは，観測値を増加させると，推定の精度が次第によくなるという性質である。大きさ n のパラメータの推定量を，$\hat{\theta}=\hat{\theta}(X_1,\ X_2,\ \cdots\cdots,\ X_n)$ とした場合，$\hat{\theta}(X_1,\ X_2,\ \cdots\cdots,\ X_n)$ が θ に確率収束する。つまり，任意の $u>0$ に対して，$P\{|\hat{\theta}(X_1,\ X_2,\ \cdots\cdots,\ X_n)-\theta|>u\}\underset{n\to\infty}{\longrightarrow}0$ を満足することをいう。一致性は，パラメータの推定値の分散が $n\to\infty$ の時，0に近づけばみたされる。そこで最小自乗推定量 b の分散を考えてみると，

$$V(b)=E\{(\Sigma u(x-\bar{x}))^2\}/(\Sigma(x-\bar{x})^2)^2$$

ここで，$E(u_i^2)=\sigma^2$ でかつ $E(u_iu_j)=0$ ゆえ，

$$V(b)=\sigma^2\Sigma(x-\bar{x})^2/(\Sigma(x-\bar{x})^2)^2$$

$$=\sigma^2/\Sigma(x-\bar{x})^2$$

n を限りなく大きくすれば $\Sigma(x-\bar{x})^2\to\infty$, したがって b の分散は限りなく0に近づく，それ故，b は β の一致推定量である。同様にして a も α の一致推定量となる。

θ の推定量，$\hat{\theta}=\hat{\theta}(X_1,\ X_2,\ \cdots\cdots,\ X_n)$ が不偏性と一致性をみたしている場合，有効性とは，$\hat{\theta}(X_1,\ X_2,\ \cdots\cdots X_n)$ と θ の誤差が他の推定量と比較して小さい可能性が大きいことを意味する。回帰分析では，最小自乗推定量が有効性をみたすように，観測値が正規分布することを仮定するのである。

上記の仮定に加えて，誤差項の正規性を仮定すれば，最小自乗推定量はさらに最適なものになる。観測値が，$(x_1,\ y_1)$，$(x_2,\ y_2)$，……，$(x_n,\ y_n)$ という値をとる確率が最大になるのが，α，β が a，b という値をとる時である場合，このような推定量のことを最尤推定量という。誤差項が正規分布する場合，尤度関数は，

$$L=(2\pi)^{-n/2}\sigma^{-n}\exp\{-\Sigma(y-\alpha-\beta x)^2/2\sigma^2\}$$

$$-\log L=n\log(2\pi)/2+n\log\sigma+\Sigma(y-\alpha-\beta x)^2/2\sigma^2$$

L が最大となるのは，$n\log(2\pi)/2$, $n\log\sigma$ が一定であるから，$\Sigma(y-\alpha-\beta x)^2$ が最小の時である。ゆえに，仮定がすべてみたされている場合には，最小自乗推定量は最尤推定量となる。

できる。また，母集団回帰線も線形である。すなわち，$E(Y|X)=A+BX$（線形性の仮定）

〔仮定2〕 従属変数 y は確率変数であり，y の母集団 Y の値は正規分布する。

〔仮定3〕 撹乱項 u は確率変数である。u はたがいに独立でありかつ x の値とも独立である。さらに $u \sim N(0, \sigma^2)$。

〔仮定4〕 撹乱項 u の標準偏差は，x の値のいかんをとわず一定である。（分散斉一性の仮定）

〔仮定5〕 回帰分析を通じて獲得された結果は，x の観測範囲内でのみ有効である。

〔仮定1〕，すなわち，線形性の仮定については，まず，分析の準備段階でスキャッター・グラフをえがくことが重要である。観測値に直線をあてはめることができるかどうかは，直観的にではあるが，この方法によってもっとも容易に把握することができる。スキャッター・グラフによって観測値に直線をあてはめることが困難である状況でも，技術的には最小自乗法によって当該観測値に対してもっとも適合する直線をあてはめることは可能であるので，この作業は不可欠なのである。

ついで，〔仮定1〕に関しては次のことに注目したい。すなわち，原価方程式獲得のための基礎データ，つまり母集団から抽出された標本（観測値）が時系列データであることである。しかるに，われわれの〔仮定4〕では，分析上では観測値は時系列データではないとみなしているのである。時系列データをもとに回帰分析を行う場合，一般に単純線形回帰モデルを適用することには問題があるとされている。(27) 線形性の仮定を設けるのは，特に獲得された原価方程式の操作性を考慮してのことである。さらに，操作性の問題のみならず，測定上の誤差を別としても，実際には厳密な直線関係が存在せずとも直線をあてはめることに意義が認められるという状況では，線形性の仮定はそれほど強い仮定ではないように思われる。

〔仮定2〕，〔仮定3〕，〔仮定4〕はそれぞれ，確率変数である y と u の母集団に関する仮定である。これらの仮定をおいてはじめて，われわれは分析を開始することができる。〔仮定2〕，〔仮定3〕は，回帰分析によって，回帰直線を推定する作業のみを考える場合には不必要なものである。しかし，信頼限界についても言及しようとするなら，

(27) 河原〔15〕p. 104.

観測値yは正規分布にしたがうと仮定しなければならない。なお，〔仮定4〕のうち，$E(u_i)=0$ はそれほどきつい仮定ではないことをのべておこう。$E(u_i)=\mu\neq0$ とすれば，

$$\begin{aligned}E(y_i)&=E(\alpha+\beta x_i+u_i)\\&=E\{(\alpha+\mu)+\beta x_i+(u_i-\mu)\}\end{aligned}$$

ここで $(\alpha+\mu)$ を定数項，$(u_i-\mu)$ を撹乱項と考えれば，$E(u_i-\mu)=0$ となるからである。このこととは直接的には関係がないが，y_i の測定精度は独立変数ほどではなくても，それほど分析結果に支障をあたえないことも一般に知られている。(28)

〔仮定4〕が成立しない状況は分散非斉一（heteroscedasticity）とよばれる。分散が非斉一の場合には，(17)式から明らかなように，回帰の標準偏差をもとめる式は無意味となる。分散が非斉一であるかどうかの検定を行うことは困難である。なぜなら，分散非斉一の場合に検定を行うには（通常，獲得不能な）比較的多量の観測値が必要だからである。(29) しかし，分散非斉一の状況は原価の動きを操業度を代表する測定尺度で説明しようとする場合には一般的である。そこでは，分散非斉一性の修正の方法を考えねばならない。

最後の〔仮定5〕は，回帰分析の結果を会計的にどのように解釈するか，つまり固定費のもつ意味にかかわる興味深い仮定である。従来，会計では獲得された回帰方程式を用いて，独立変数が0の値をとるものとした場合（通常，0操業時の観測値は獲得不能であるが）にえられる従属変数 y の推定値を固定費と解釈して使用することが多かった。回帰分析の分析対象である原価が費目原価であるか部門原価であるか，それとも総原価であるかによって，考慮すべき事項には若干の相違がある。しかし，観測値に0操業時のデータが存在しないかぎり，0操業を回帰方程式上で想定し，原価推定値をえるという外挿法の適用は厳にいましめられるべきである。この意味でわれわれは〔仮定5〕を，回帰分析の誤った使用がなされてはならないという警告としてうけとらねばならない。したがって，従来議論されてきた上記のような手続でえられた推定値が負の値をとろう

(28)　同様の記述が次の文献にもみられる。
Benston [1] p. 665, *fn.* 15., Bierman and Dyckman [2] p. 68.

(29)　「この検定は観測値ごとにその実際値の分散を $\hat{y}_i$ のまわりで計算することにより行われる。この際，十分な数の観測値が存在しなければ分散を算出してそれらを比較することはできない。」Dopuch *et. al.* [4] pp. 78—79.

とも，たとえ正の値であっても，このような操作はすべて無意味である。なお，後節でのべる重回帰分析の場合には，独立変数の数が複数となり，観測値の存在領域が多次元空間であらわされることになるので，外挿法が適用される危険性が増大するから特に注意すべきである。

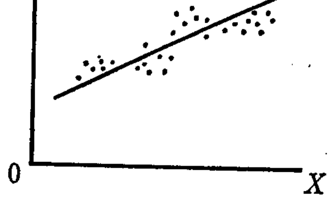
〔仮定3〕のうちで仮定された撹乱項相互の独立性が侵犯される状況が実際には存在するだろう。原価と独立変数間で，実際には線形ではあらわせない関係が存在するにもかかわらず，線形関数をあてはめた場合，独立変数値が大きくなれば，撹乱項もそれに応じて増加する傾向がある(30)場合などがこの状況にあてはまる。このように撹乱項が互いに独立でない状況は自己相関（autocorrelation）あるいは系列相関（serial correlation）とよばれる。自己相関が存在する場合，スキャッター・グラフに観測値をプロットし，回帰直線をそれに重ねると，データは回帰直線を中心として波形にうねっているようにあらわれる(31)。「回帰分析が，季節的変動（seasonality phenomena）が存在するかどうか不明瞭な原価に対して適用された場合，観測値がただ単に自己相関の影響をうけているだけなのに，あたかも季節的変動を呈示しているかのように誤解される場合がある(32)」ので注意すべきである。

また，自己相関が存在する場合には，標準誤差と有意性判定の公式は，近似式としても適用できなくなるので，誤差項（母集団回帰直線からの観測値 y の偏差をいう）が実際に，確率変数であるかどうかを確認することは非常に重要な問題である。一般に，誤差項 ε_i の自己相関は，残差項 u_i に反映されると考えられる。そこで自己相関の検定には，ダービン＝ワトソン比（ Durbin-Watson ratio）; d が使用される。

---

(30)　Dopuch *et. al.* [4] p. 80.

(31)　下図のような状況が自己相関があらわれていることを示している。



このように，スキャッター・グラフは自己相関の有無を確認するのにも使用できる。

(32)　Corcoran [3] p. 62.

u_t を期間 t の残差項，つまり，

$$u_t=\widehat{y}_t-y_t \quad (t=1,\ 2,\ 3,\ \cdots\cdots,\ m)$$

としたとき，

$$d=\sum_{t=2}^{m}(u_t-u_{t-1})^2/\sum_{t=2}^{m}u_t^2 \tag{19}$$

一般に，確率変数 x_i で $E(x_i)=0$, $E(x_i^2)=\sigma^2$ なら，$E(x_i-x_j)^2=2\sigma^2(i\neq j)$ である。したがって，残差項が相互に独立なら，d の値は2であり，正の自己相関がある場合には2より小さくなる。負の自己相関が存在することはまれであるが，その場合には，$4\geq d>2$ となる。

y_t の値が y_{t-1} の値と独立でなければ，残差項 u_t も残差項 u_{t-1} とは独立ではない。つまり，$u_t=pu_{t-1}+$誤差（ここで $|p|<1$）とした時に，$p=0$ なら，y の値の独立性が明らかになり，残差項もたがいに独立となる。そこで，帰無仮説（null hypothesis）H_0：$p=0$，対立仮説（alternative hypothesis）H_1：$p>0$ を設定し，ダービン＝ワトソン比の下限の有意点 d_L，上限 d_U，と(19)式で算定される d の値を比較し，

$$\begin{cases} d\geq d_U \text{ なら，} H_0 \text{ は棄却できない} \\ d_L<d<d_U \text{ なら，検定は不能} \\ d\leq d_L \text{ なら，} H_0 \text{ は棄却される} \end{cases}$$

という結論を導くのである。(33)

以上の検定手順によって，自己相関の存在の有無を決定する。$d\leq d_L$ の時には，後述の方法によっても自己相関が解消されなければ，時系列データを最小自乗法が適用できるように修正する必要があろう。この修正が行われなければ，回帰係数の信頼限界をもとめることは適当ではない。また，ダービン＝ワトソン比を使用する時，信頼するにたる結果を獲得するためには，標本の大きさが50以上必要であることにも注意すべきである。(34) 自己相関は，モデルに新たな変数を導入したり，変数変換を行ったり，あるいはダミー変数（dummy variable）を導入することによって解消されることもある。(35) 上記のダービン＝ワトソン比を用いた検定で結果が不能な場合には，試みられてしかるべきである。

(33)　5％有意水準での d_U と d_L の値を表にしたのが表2である。なお，n は観測値数，k' は独立変数の数である。表2は，Durbin and Watson [5] p. 173 から引用した。

(34)　Corcoran [3] p. 62.

表2　5％有意水準での d_L, d_U

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

以上，回帰分析の適用にあたって考慮しなければならない諸仮定について説明を加えてきた。次節では，それらのいくつかについてさらに詳細に検討を加えていくことにす

(35) *Ibid.*, pp. 62—63.

たとえば，変数変換については佐和・前川〔19〕pp. 31—32，ダミー変数の利用については，Benston [1] pp. 666—667. を参照せよ。

る。

3－6　分散共分散マトリクスの利用と信頼区間の公式

節3－4までの議論では，観測値はすべて，その母集団から抽出された標本であることを明示せずに説明を行ってきた。一般回帰分析では，パラメータの母集団は $\boldsymbol{\beta}$ であり，観測値 y は次式でしめされる母集団から抽出された標本である。

$$\boldsymbol{Y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon} \quad (\boldsymbol{\varepsilon}\text{は誤差項の } n\times 1 \text{ ベクトル}) \tag{20}$$

誤差項は前述した残差項と密接に関連するが同一のものではない。つまり，誤差項が母集団回帰直線からの観測値の偏差であるのに対して，残差項は推定回帰直線からの観測値 y の偏差なのである。誤差項と残差項の混乱を防ぐため，以下では誤差項を ε であらわすことにする。誤差項 ε_t に関しては次のことが仮定されている。

(1)　ε_t は平均0，標準偏差 σ の正規確率変数である。すなわち，$E(\varepsilon_t)=0$, $V(\varepsilon_t)=\sigma^2$ かつ，$\varepsilon_t\sim N(0,\sigma^2)$

(2)　ε_s は ε_t とは相互に独立（$s\neq t$）で，かつ $\varepsilon_s\sim N(0,\sigma^2)$

(3)　$E(\varepsilon_t\varepsilon_s)=0$ ただし，$s\neq t$，つまり，誤差項は無相関である。(36)

誤差項の分散が σ^2 であることをマトリクス表示すれば次のようになる。

$$E(\boldsymbol{\varepsilon\varepsilon}^T)=\sigma^2\boldsymbol{I} \tag{21}$$

(1), (2)から明らかなように，独立変数の値のいかんにかかわらず，それぞれの誤差項の分散が等しい（$=\sigma^2$）ことは前述の分散斉一性を意味する。σ は未知であるので，推定に際してはわれわれはその代用物として，モデルが正しいと仮定した上で，標本分散（sampling variance）；$\hat{S}^2_{y\cdot\cdot x}$ を使用し，近似計算を行ってきたのである。

パラメータの母集団値 $\boldsymbol{\beta}$ とその推定値 $\hat{\boldsymbol{\beta}}$ の間には次のような関係が存在する。

$$\begin{aligned}\hat{\boldsymbol{\beta}}&=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}\\&=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T(\boldsymbol{X\beta}+\boldsymbol{\varepsilon}) \qquad (\because\ (20)\text{式より})\\&=(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{\beta}+(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon}\\&=\boldsymbol{\beta}+(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon}\end{aligned} \tag{22}$$

推定値 $\hat{\boldsymbol{\beta}}$ の分散は次のようになる。

$$V(\hat{\boldsymbol{\beta}})=\boldsymbol{E}[(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T]$$

---

(36)　$E(\varepsilon_s\varepsilon_t)\neq 0$ の場合，自己相関の状況が呈示される。Corcoran [3] p. 64.

$$=E[(\boldsymbol{\beta}-(\boldsymbol{\beta}+(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon}))(\boldsymbol{\beta}-(\boldsymbol{\beta}+(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon}))^T]$$
$$=E[((\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon})((\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon})^T]$$
$$=E[(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}]$$
$$=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T)\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}$$
$$=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\sigma^2\boldsymbol{I}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (\because \text{(21)式より})$$
$$=(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{X})\sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (\because \ \sigma \text{はスカラー})$$
$$=\sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1} \text{(37)} \tag{23}$$

σ^2は未知なので，その推定量$\widehat{S}^2_{y\cdot x}$を使用すると，

$$V(\widehat{\boldsymbol{\beta}})=\widehat{S}^2_{y\cdot x}(\boldsymbol{X}^T\boldsymbol{X})^{-1} \tag{24}$$

(24)式の結果を分散共分散マトリクス（variance-covariance matrix）といい，以下では$\boldsymbol{V}$であらわすことにする。

$$\boldsymbol{V}=\widehat{S}^2_{y\cdot x}(\boldsymbol{X}^T\boldsymbol{X})^{-1}=\text{分散共分散マトリクス}^{(38)} \tag{25}$$

$\boldsymbol{X}$の行ベクトルを$\boldsymbol{x}_i(=[1\ x_i])$とすると，$\widehat{y}=\boldsymbol{x}_i\widehat{\boldsymbol{\beta}}$であるから，推定値$\widehat{y}$の分散は，

$$V(\widehat{y})=V(\boldsymbol{x}_i\widehat{\boldsymbol{\beta}})=\boldsymbol{x}_i\boldsymbol{V}\boldsymbol{x}_i^T \tag{26}$$

(26)式は次のようにもあらわされる。

$$V(\widehat{y})=\widehat{S}^2_{y\cdot x}\left[\frac{1}{n}+\frac{(x-\bar{x})^2}{\Sigma(x-\bar{x})^2}\right] \tag{27}$$

それでは，分散共分散マトリクスがいかに有用であるかをしめすことにしよう。表1のデータから(25)式であらわされる分散共分散マトリクスをもとめると次のようになる。

$$\boldsymbol{V}=\widehat{S}^2_{y\cdot x}(\boldsymbol{X}^T\boldsymbol{X})^{-1}=\begin{bmatrix}730.569 & -9.8667\\ -9.8667 & 0.129399\end{bmatrix}$$

(25)式から，たとえば $x=55$ とすれば，

$$V(\widehat{y})=[1\quad 55]\boldsymbol{V}\begin{bmatrix}1\\55\end{bmatrix}=86.66$$

(37)　(23)式は次のように書きあらわせる。

$$V(\widehat{\boldsymbol{\beta}})=\frac{\sigma^2}{n\Sigma(x-\bar{x})^2}\begin{bmatrix}\Sigma x^2 & -\Sigma x\\ -\Sigma x & n\end{bmatrix}$$

(38)　分散共分散マトリクスの要素は以下にしめすとおり。

$$\boldsymbol{V}=\begin{bmatrix}V(a) & \mathrm{Cov}(a,\ b)\\ \mathrm{Cov}(a,\ b) & V(b)\end{bmatrix}$$

したがって，$x=55$ の時の信頼限界は次のようになる。

$$\hat{y}\pm t_{n-2,\alpha/2}\sqrt{V(\hat{y})}=56.886+3.25\times 55\pm 2.228\sqrt{86.66}$$

係数$\boldsymbol{\beta}$（a と b）のそれぞれの標準偏差は，分散共分散マトリクスの主対角要素の平方根をとればもとめられる。固定費（a）の標準偏差は $\sqrt{780.569}=27.94$（単位1,000ドル），変動費率（b）の標準偏差は，$\sqrt{0.129399}=0.36$ である。当然のことであるが，係数の標準偏差はその値が小さければ小さいほど，係数の信頼性が増大する。

操業度が1単位変動した場合の限界原価の値は，$\hat{y}$ の第1次導関数をとればもとめられる。

$$\frac{d\hat{y}}{dx}=b=3.25$$

限界原価の95%信頼限界は次のようになる。

$$3.25\pm 2.228\times 0.36=\{2.44792,\ 4.05208\}$$ [39]

この場合にでも，サンプルサイズが大きければ（$n\geq 30$），サンプルが正規分布するとみなして信頼限界を算出することができる。

「標準誤差（本稿でいう回帰の標準偏差のこと……筆者注）によって推定値の信頼性を測定しようとする場合，……，この標本は母集団から抽出された1つの標本にすぎず，多様な標本を同じ母集団から抽出できるのであって，これらの異なる標本を用いるとき，同じ回帰線がえられるとは考えられないので，回帰分析に標本誤差を避けることはできない。そこで，推定値 a，b に関する標準誤差の測定の問題とこれらの推定値にもとづいてなされる統計的推論を考察する必要がある。[40]」分散共分散マトリクスはまた，後者の回帰分析の結果の検定にも使用できる。回帰直線が母集団に適合しない場合には，母集団回帰線 $E(Y|X)=A+BX$ の傾きは0（すなわち，$B=0$）となると考えられる。帰無仮説 $H_0: B=0$ と対立仮説 $H_1: B\neq 0$ を設定する。95%信頼水準のもとで，自由度 $n-2=10$ の t 値は ± 2.228 であり，これと次式でもとめられる検定統計値tとを比較

---

(39)　y の個々の推定値についての信頼限界は，

$$\hat{y}\pm t_{n-2,\alpha/2}\sqrt{V(\hat{y})-\hat{S}^2_{y\cdot x}}$$

この信頼限界に関する理論的解釈については注24をみよ。Corcoran [3] p. 66, *fn.* 4.

(40)　豊島〔21〕pp. 135—136.

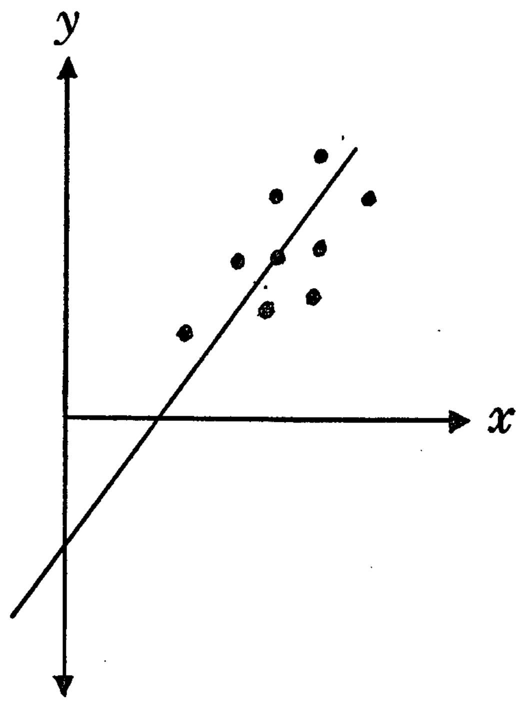
するのである。

$$t=\frac{b-B}{\sqrt{V(b)}}=\frac{3.2474-0}{0.36}=9.02 \tag{28}$$

この場合，t 値が規準値2.228をこえるので，帰無仮説は棄却され，X と Y の間に線形関係は存在しないという考えは棄てざるをえない。[41]

同様の方法で，y 切片の t 検定を行うことは技術的には可能である。しかしながら，このような検定は，観測値に y 軸上の数値が含まれる場合（通常，このような観測値は存在しない）をのぞいて実質的には無意味である。つまり，y 切片のt検定は，回帰分析への外挿法の適用にほかならないのである。特に図4のように回帰直線の y 切片がマイナスの値をとる時には，回帰直線は観測範囲内の原価額を予測することのみに有効である。統計学で外挿法の適用を強くいましめることは，会計における有効操業圏（relevant range）の概念と相通じるところがある。回帰直線の y 切片がマイナスの値をとることは，本来，固定費の範疇に属する原価項目を会計的な種々の制約によって変動費のごとく処理する結果，変動費率が過大評価されたと解釈できなくはない。したがって，このような状況は，コーコランの指摘するように，会計データをまずラフに固定費，変動費，および準変動費に分解し，しかるのち，回帰分析を変動費および準変動費に対して適用することが回帰分析を行うにあたって有意義であることを示唆するものである。[42]

図 4



(41) サンプルサイズが30をこえる時，当該サンプルの母集団は正規分布するものとして，統計量 $Z=(b-B)/\sqrt{V(b)}$ が検定に使用できる。Corcoran [3] p. 67, *fn.* 6.

(42) *Ibid.*, p. 67. なお，コスト・ビヘイビアの把握の上で重要となる検討方針についての記述がコーコランの接近方法とは若干異なるが，豊島〔21〕pp. 161—121. にある。

3-7　分散分析とF検定

回帰直線の検定には前節で説明したt検定以外にF検定がある。t検定は係数ごとに原価方程式を検定する方法であるが，F検定はF分布の特性を利用して係数全体で推定式のすべてを検定する方法である。

回帰方程式が母集団に適さない場合には，傾き（回帰係数）は0になると考えられる。つまり，

$$E(b)=\beta_1=0 \quad (\beta_1 \text{は} b \text{の母集団値})$$

しかし，回帰方程式が母集団の様子を的確にあらわしていればいるほど，誤差変動が減少し，相対的に回帰変動が増加する。誤差変動，回帰変動をそれぞれの自由度で除した分散の値は平均平方（mean square；不偏分散ともいわれる）とよればる母集団分散の推定値である。誤差変動の不偏分散で回帰変動のそれを除した値がF値（F-ratio）とよばれるものである。F値からわれわれは次にしめす事項を知ることができる。

(1)　F値は，誤差変動の分散推定値の何倍くらい回帰変動の分散推定値があるかを測定する。

(2)　F値が上昇する可能性があるかどうかの決定を行うことを可能にする。(43)

通常，F値の計算を容易にするために分散分析表（analysis of variance table；

表3　分散分析表 (ANOVA)

変動要因	自由度	平方和	不偏分散
回帰変動(R)	m	$SSR=\Sigma(\bar{y}-\hat{y})^2$	$MSR=SSR/m$
誤差変動(E)	$n-m-1$	$SSE=\Sigma(\hat{y}-y)^2$	$MSE=SSE/(n-m-1)$
全変動(T)	$n-1$	$SST=\Sigma(\bar{y}-y)^2$	$MST=SST/(n-1)$

m；独立変数の数　n；観測値数

表4　分散分析表

変動要因	自由度	平方和	不偏分散
回帰変動	1	27,611.00	27,611.00
誤差変動	10	3,388.00	338.80
全変動	11	30,999.00	2,818.09

(43)　Corcoran [3] p. 68.

ANOVA）が作成される。表3がそのひな型である。表4は表1のデータをもとに作成した分散分析表である。F 値を計算すると，

$$F=\frac{MSR}{MSE}=\frac{27,611}{338.8}=81.50 \tag{29}$$

5％棄却域を設定する場合のFの規準値は，

$$F_{1,10}=4.96$$

である。われわれの計算例では，$F>F_{1,10}$ なので帰無仮説 $H_0:\beta_1=0$ は棄却され，観測値は母集団から抽出されたものと結論される。このように，F 検定は全体的に回帰直線のあてはまりのよさ（goodness of fit）を測定するのである。F 検定は，たとえば，「操業度は5％棄却域で有意であった」とか「予測方程式には95％の信頼性がある」とかというような確率的状況について言明できるという点で，ただ単に変数間の関連の度合を測定する統計値である決定係数よりもすぐれたものである。

3－8　回帰分析のコンピュータプログラム

前節までで説明してきた回帰分析を実行するコンピュータプログラムをコーコランはしめしている。(44) これは高級言語 BASIC（Beginner's All-purpose Symbolic Instruction Code）で書かれたもので，彼自身が指摘・強調するように，(1)独立変数が2つ以上の場合の回帰分析（次節でのべる重回帰分析）にも適用でき，また，(2)必要ありと認められた場合には種々の関数でもって回帰分析を行うこともできる。後者の場合には変数変換を行って，(45) 当該関数を線形に変換してから，$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$ が使用できるように仕組

(44) *Ibid.*, pp. 70—74.

(45) 例えば，次のような処理をほどこすことにより，線形変換が可能となる。

	関数	データの修正
1	$y=b_0+b_1x+b_2x^2$	x，x^2 を x_1，x_2 として処理
2	$y=ax^b$	$\log y=\log a+b\log x$
3	$y=ae^{bx}$	$\ln y=\ln a+bx$
4	$y=a+\frac{b}{x}$	$y=a+b\left(\frac{1}{x}\right)$
5	$y=\frac{1}{a+bx}$	$\left(\frac{1}{y}\right)=a+bx$

Ibid., p. 70.

まれている。[46,47]「BASICではサブルーチンと主プログラムとは直接関連していて，サブルーチンの独立性はほとんどない」[48]ためその一部分をとりだして説明することは非常に困難である。そこで，本節までの記述を理解した上で，彼のプログラムをいくつかの参考文献[49]を使って検討していただきたい。BASICは非常に理解が容易な科学計算用の言語であり，FORTRANを学習したものであれば短時間で理解できるものである。ただ，本プログラムを実行する際には，言語にコンパイラ間で著しい相違があるので注意すべきである。

3－9　重回帰分析

「重回帰分析は，複数の独立変数がある場合に単純回帰分析の概念を拡張したものである。」[50]重回帰分析の概要説明をまつまでもなく，独立変数が相互に密接に相関する（例えば，機械運転時間と修繕時間を独立変数に採用する場合を考えてみよう。通常の場合，実際に両者の相関の度合の算定結果をまつまでもなく，このような関係が存在することは容易に予想することができる）という事態が発生する可能性がある。このような状況は重共線性（multicolinearity）とよばれる。数多くの独立変数を回帰分析に導入することは技術的には可能であるが，[51]変数の数がますにしたがって，必然的に重共線性が問題になる。したがって，分析に3つ以上の独立変数を導入することは無意味である。[52,53]また，

---

(46)　*Ibid.*, p. 70.

(47)　線形変換のプログラムは，当該プログラムに追加される必要がある。

(48)　田中〔20〕p. 9.

(49)　次のものを参照されたい。

Kemeny, J. G. and T. E. Kurtz, *Basic Programming*, John Wiley and Sons, Inc., 1967.（森口繁一監訳，尾崎義雄・神山武訳「ベーシック入門」，共立出版，1971.），味村重臣・岡田雍彦・渡辺美枝著「入門BASIC」，オーム社，1971.，田中〔20〕.

(50)　Corcoran [3] p. 75.

(51)　Horngren [6] p. 832.

(52)　Corcoran [3] p. 76. コーコランは，統計学者 M. J. Moroney の主張に同意するものであるとして次の文献を引用している。

Moroney, M. J., *Facts from Figures*, Penguin Books, 1967, p. 304. *Ibid.*, p. 76, *fn.* 7.

(53)　「計量分析の目的は予測であり，したがって，できるだけ2変数回帰を用いるべ↗

重共線性が存在するときには，回帰係数の分散共分散マトリクスの各要素の値が大きくなり，標準偏差の値はゆがめられ，有意性の検定（t 検定）は実質的には行えなくなる。

独立変数の数が増しても y および $\bar{y}$ の値は不変であるから，全変動，回帰変動，誤差変動の数値は単純回帰分析の場合と同じである。$\hat{y}$ の値は，$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$ からえられるパラメータの推定値をもちいて算定できる。単純回帰分析の時と同様に最小自乗法を適用し，撹乱項の平方和をとり，それを最小化することにより，2変数回帰の場合の正規方程式が獲得できる。

$$\text{最小化}\quad E=\Sigma u_i^2=\Sigma(y-\hat{y})^2 =\Sigma(y_i-b_0-b_1x_{i1}-b_2x_{i2})^2 \qquad (30)$$

$$(\hat{y}_i=b_0+b_1x_{i1}+b_2x_{i2})$$

$$\frac{\partial E}{\partial b_0}=\frac{\partial E}{\partial b_1}=\frac{\partial E}{\partial b_2}=0 \qquad (31)$$

$$\begin{cases} \Sigma 2(y_i-b_0-b_1x_{i1}-b_2x_{i2})(-1)=0 \\ \Sigma 2(y_i-b_0-b_1x_{i1}-b_2x_{i2})(-x_{i1})=0 \\ \Sigma 2(y_i-b_0-b_1x_{i1}-b_2x_{i2})(-x_{i2})=0 \end{cases} \qquad (32)$$

(32)式を整理して，

$$\begin{cases} b_0n+b_1\Sigma x_1+b_2\Sigma x_2=\Sigma y \\ b_0\Sigma x_1+b_1\Sigma x_1^2+b_2\Sigma x_1x_2=\Sigma x_1y \\ b_0\Sigma x_2+b_1\Sigma x_1x_2+b_2\Sigma x_2^2=\Sigma x_2y \end{cases} \qquad (33)$$

(54)

（添字は省略）

きである，と主張する人々もいる。」佐和・前川訳〔20〕p. 47. レッサーは，かかる主張の論客として Geary をあげている。Geary, R. C., Some Results about Relations between Stochastic Variables; A Discussion Document, *Review of International Statistical Institute*, 31, 1963, pp. 163—181.

(54) 観測値数を n，独立変数を $x_i(i=1, 2, \cdots\cdots, k)$，回帰係数を $b_i(i=1, 2, \cdots\cdots, k)$，定数項を b_0 とすれば，正規方程式の一般形は，

$$\begin{cases} b_0n+b_1\Sigma x_1+\cdots\cdots+b_k\Sigma x_k=\Sigma y \\ b_0\Sigma x_1+b_2\Sigma x_1^2+\cdots\cdots+b_k\Sigma x_1x_k=\Sigma x_1y \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \qquad\quad \vdots \\ b_0\Sigma x_k+b_2\Sigma x_1x_k+\cdots\cdots+b_k\Sigma x_k^2=\Sigma x_ky \end{cases}$$

となる。

それでは重回帰分析についての考察を続けるために表1に表5のデータを追加する。

表　5

年度	y 原価	x_1 機械運転時間	x_2 直接作業時間	x_3 卸売物価指数
1963	262	60	16	94.5
1964	230	55	13	94.7
1965	247	72	15	96.6
1966	258	62	15	99.8
1967	240	58	14	100.0
1968	330	75	21	102.5
1969	314	74	16	106.5
1970	340	85	18	110.4
1971	348	88	19	113.9
1972	360	94	18	119.1
1973	350	92	15	134.7
1974	375	100	14	160.1

（単位　1,000）

表5をみてもわかるように，新しく2つの独立変数——直接作業時間と卸売物価指数(Wholesale Price Index；以下WPIと略称する）——が導入されている。

さて，3つの独立変数をすべて分析に組込む場合，マトリクスXは次のようになる。

$$X = \begin{pmatrix} 1 & 60 & 16 & 94.5 \\ 1 & 55 & 13 & 94.7 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 100 & 14 & 160.1 \end{pmatrix}$$

コーゴランのしめしたコンピュータプログラムにかけると，推定方程式ならびに決定係数およびF値は次のようになる。

$$\hat{y} = -100.82 + 1.309x_1 + 9.570x_2 + 1.357x_3 \tag{34}$$

$$r^2 = 0.964$$

$$F = 71.5$$

分散共分散マトリクスは下記のようになる。

$$V=\begin{pmatrix} 2,062.67 & 20.80 & -96.36 & -18.72 \\ 20.80 & 0.46 & -1.25 & -0.32 \\ -96.36 & -1.25 & 5.68 & 0.90 \\ -18.72 & -0.32 & 0.90 & 0.26 \end{pmatrix}$$

前述の決定係数の値とF値は(34)式の回帰方程式を維持使用すべきことを数値上はしめしている。「ただそれらの値があまりにも好ましすぎるということから重共線性の問題をさけてとおることはできない(55)」のである。

それでは，重共線性の検定の問題を考えてみよう。「重共線性を検定するためには，各回帰係数 b_i に対する各々の回帰係数の標準誤差をチェックする必要がある。この標準誤差が各係数の値をこえる場合には，重共線性に問題があるといえる(56)。」

次のような相関係数マトリクスが重共線性をチェックするためにしばしば使用される。

$$\begin{array}{cc} & \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} & \left|\begin{array}{ccc} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{array}\right| \end{array} = \begin{array}{cc} & \begin{array}{ccc} \text{機械運転時間} & \text{直接作業時間} & \text{WPI} \\ x_1 & x_2 & x_3 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} & \left|\begin{array}{ccc} 1 & 0.3422 & 0.8419 \\ 0.3422 & 1 & 0.0878 \\ 0.8419 & 0.0878 & 1 \end{array}\right| \end{array}$$

各々の相関係数は次のように計算される。

$$r_{ij}=\frac{\mathrm{Cov}(x_i,\ x_j)}{\sqrt{V(x_i)}\sqrt{V(x_j)}}=\frac{[n\Sigma x_i x_j-\Sigma x_i \Sigma x_j]/n^2}{\sqrt{(n\Sigma x_i^2-(\Sigma x_i)^2)(n\Sigma x_j^2-(\Sigma x_j)^2)/n^2}} \tag{33}$$

機械運転時間とWPIの相関を考えると，マトリクス X^TX は次のようになる。

$$\begin{pmatrix} 12 & 915 & 194 & 1,332.8 \\ 915 & 72,387 & 14,930 & 104,394.8 \\ 194 & 14,930 & 3,198 & 21,502.6 \\ 1,332.8 & 104,394.8 & 21,502.6 & 152,160.72 \end{pmatrix}$$

上記のマトリクスは次のことを示している。

$e_{12}=\Sigma x_1=915$　　$e_{22}=\Sigma x_1^2=72,387$

$e_{14}=\Sigma x_3=1,332.8$　　$e_{44}=\Sigma x_3^2=152,160.72$

(55) Corcoran [3] p. 76.

(56) *Ibid.*, p. 77.

$e_{24}=\Sigma x_1x_3=104,394.8$

WPI に対して機械運転時間を回帰させた場合の決定係数の平方根は，

$$r_{13}=\frac{n\Sigma x_1x_3-\Sigma x_1\Sigma x_3}{\sqrt{(n\Sigma x_1{}^2-(\Sigma x_1)^2)(n\Sigma x_3{}^2-(\Sigma x_3)^2)}}$$

$$=\frac{ne_{24}-e_{12}e_{14}}{\sqrt{(ne_{22}-e_{12}{}^2)(ne_{44}-e_{14}{}^2)}}$$

$$=\frac{12\times104,394.8-915\times1,332.8}{\sqrt{(12\times72,387-915^2)(12\times152,160.72-1,332.8^2)}}$$

$$=0.841889$$

である。

「このように r の値が大きいことは，機械運転時間か WPI のどちらか一方を分析に組込む独立変数から除外することを示唆するものである。[57]」ここで，コーコランは，回帰推定値の利用目的に応じて，独立変数の選択をかえねばならぬことを主張している。すなわち，(1)回帰線から遠くへだたった，(2)相互に相関する変数を用いて原価予測を行わないかぎり，重共線性の問題は回帰方程式を原価予測に用いる場合にはこれを無視してもさしつかえないだろう。一方，他の独立変数の値を一定として1つの独立変数の単位当たりの変動（つまり，限界原価）を推定しようとする場合には重共線性は重大な問題となる。すなわち，重共線性が存在するときには，ある独立変数の値は他の独立変数値の影響をうけるので $\partial y/\partial x_i$ の値を一定にしておくことはできない。したがって，他の条件が等しければという仮定に抵触するので限界原価の値は一意的には定まらないのである。[58] われわれの例で，相関係数マトリクスから重共線性がもっとも低いのは，作業時間と WPI の間である。作業時間と WPI を独立変数，原価を従属変数として回帰分析を行うと，回帰方程式，r^2，および F 値は次のようになる。

$$\hat{y}=-160.56+13.148x_2+2.273x_3$$

$$r^2=0.947$$

$$F=80.6$$

次年度の直接作業時間が 15,000 時間，WPI が 172 であるとすれば，当該年度の原価推定額は次のようになる。

---

(57)　*Ibid.*, p. 78.

(58)　*Ibid.*, p. 78.

$$\hat{y} = -160,560 + 13,148 \times 15 + 2,273 \times 172$$
$$= 427,616$$

Ⅳ　コーコランの論述に対する若干のコメント――結びにかえて

以上，コーコランの論述にそって説明を加えてきた。彼の主張の特徴を要約すると次のようになる。

(1)　コスト・ビヘイビアを明らかにし原価額を把握する技法として回帰分析という統計的方法がすぐれていることを指摘している。

(2)　最小自乗推定量がすぐれた統計的特性を維持するために必要な諸仮定を明らかにし，それらが維持されない状況について考察を行っている。

(3)　統計的方法により，パラメータの推定だけではなく，従来会計ではあまり議論されることがなかった，それらの信頼性を測定しようという検定も行えることに言及している。

(4)　固定費の解釈や分析の適用状況など会計に対する統計的方法のインパクトをも考慮している。

(5)　重回帰分析の適用についても，結果としてえられる推定値の利用目的を考慮して独立変数の数を少なくおさえることを強調する。

(6)　行列演算で回帰分析を行うことがすぐれた方法であることを指摘し，分析自体をコンピュータにのせて原価分析を迅速かつ簡易に行うことを意図している。

なお，彼の特徴を明確にするために筆者が補筆したのは次の事項である。

(1)　回帰分析で未知のパラメータを推定する方法である最小自乗法によって獲得した推定量が統計的にすぐれたものであることを指摘する。

(2)　回帰分析の諸仮定について，それらを会計的側面から検討したこと。

以上のように要約される彼の論述の基礎には次にしめすような思考が存在することを忘れてはならない。すなわち，節3－1でのべたように，回帰分析は独立変数の変動によって，従属変数がどの程度説明できるかの言明を行うものである。したがって，従属変数の変動は，独立変数の変動によるものであるという変数間の因果関係は，これをしめすことはできないのである。われわれの目的は，まさに後者の関係をつかんだ上で獲得された原価方程式を有効に利用することであるのだから，原価方程式の操作上の問題をも考慮して，適切な独立変数を選択する必要がある。再三のべてきたが，独立変数の

選択の良否いかんが分析の価値を決定するといっても過言ではない。この独立変数の選択問題こそ会計の問題である。つまり，回帰分析の原価推定問題への適用は，推定の精度を確率論的に言明できる等の利点をもつことにとどまらず，分析を実施する上で，変数間の因果関係をもっぱら会計の観点から組込むことができるのである。かかる意味でコーコランの原価推定に関する記述はわれわれに非常な示唆をあたえるものであるといえる。

ただ，最後にふれておきたいのは，回帰分析の有用性が一般に明らかであるのにもかかわらず，会計においてはその具体的適用がほとんど考察されていないことである。(59) 回帰分析が原価の分析に使用された場合，原価計算システムにあたえるインパクトも多大であるばかりではなく，(60) 会計的なエレメントが原価とその影響要因間の本来的な関係をさまたげていることのシグナルをも獲得しうるのである。したがって，われわれは回帰分析の適用上の問題を今後とも検討していかねばならないのである。

参　考　文　献

〔1〕 Benston, G. J., Multiple Regression Analysis of Cost Behavior, *The Accounting Review*, Oct., 1966.

〔2〕 Bierman, H. Jr. and T. R. Dyckman, *Managerial Cost Accounting*, The Macmillan Company, 1971.

〔3〕 Corcoran, A. W., *Costs; Accounting, Analysis, and Control*, John Wiley and Sons, Inc., 1978.

〔4〕 Dopuch, N., J. G. Birnberg, and J. Demski, *Cost Accounting; Accounting Data for Management's Decisions*, 2nd ed., Harcourt Brace Jovanovich, Inc., 1974.

〔5〕 Durbin, J. and G. S. Watson, Testing for Serial Correlation in Least Square Regression, *Biometrika*, 37, pp. 409—428, 1950; 38, pp. 159—178, 1951.

〔6〕 Horngren, C. T., *Cost Accounting; A Managerial Emphasis*, 4th ed., Prentice Hall, Inc., 1977.

〔7〕 Jensen, R., Multiple Regression Models for Cost Control——Assumptions and Limitations, *The Accounting Review*, Apr., 1967.

(59) Benston [1], Jensen [7], Koehler and Neyhart [8], McClenon [9], Raun [11], 門田〔16〕, 岡本〔18〕, 豊島〔21〕を参照せよ。

(60) 岡本〔18〕p. 44 を参照せよ。

〔8〕 Koehler, R. W. and C. A. Neyhart, Difficulties in Flexible Budgeting, *Managerial Planning*, May-Jun., 1972.

〔9〕 McClenon, P. R., Cost Finding Through Multiple Correlation Analysis, *The Accounting Review*, Jul., 1963.

〔10〕 Moore, C. L. and R. K. Jeadicke, *Managerial Accounting*, 4th ed., South-Western Publishing Company, 1976.

〔11〕 Raun, D. L., The Limitations of Profit Graphs, Breakeven Analysis, and Budgets, *The Accounting Review*, Oct., 1964.

〔12〕 藤澤袈裟利，松行康夫著「経営数学」——経営学全書34巻，丸善，昭和50年。

〔13〕 今川正著「新しい経済統計学」，春秋社，昭和47年。

〔14〕 岩田暁一著「経済分析のための統計的方法」東洋経済新報社，昭和47年。

〔15〕 河原裕介著「需要予測の実際」東洋経済社，昭和53年。

〔16〕 門田安弘稿，「原価予測に対する一考察」(未発表論文)。

〔17〕 N.R.ドレーパー，H.スミス著，中村慶一訳「応用回帰分析」森北出版，昭和44年。

〔18〕 岡本清，「原価分解をめぐる諸問題」『産業経理』昭和53年10月号。

〔19〕 C.E.V.レッサー著，佐和隆光，前川功訳「初等計量経済学」東洋経済新報社，昭和53年。

〔20〕 田中良久著「BASIC入門——行動科学のためのコンピュータ・プログラミング入門」東京大学出版会，昭和50年。

〔21〕 豊島義一著「意思決定原価計算」同文館，昭和54年。