HITAC基本統計システムにおける重回帰プログラムとデータスクリーニング

沢　　勲　*Isao SAWA*
采山寛幸　*Hiroyuki UNEYAMA*
安　　光烈　*Guanglie AN*

ABSTRACT

Computer processing has been done of the data from the tests of physique and physical fitness of students at Osaka University of Economics and Law.

We got the data from ten kinds of test statistically processed. Physique examination involves that of height, weight and chest : Physical fitness test covers that of side step, vertical jump, back strength, grip strength, trunk extension, standing trunk flexion as well as step test.

We took adoped for data of sports test on screening method. The information processing includes the following items : mean value, standard error, regression coefficient, F-value in analysis of variance for the regression and computed t-value of multiple regression statistics.

1　はじめに

文部省では、1964年以来、体力・運動能力調査（スポーツテスト、壮年体力テスト等）を実施している。それは、国民の体力および運動能力の現状を明らかにするためである。その結果、大学生の健康・体力づくりに資し、体育・ス

ポーツ活動の指導や行政上の資料としている。文部省の資料では[1]、統計数値表をはじめ、体力・運動能力の年齢別推移、運動・スポーツの実施状況別体力・運動能力の比較がある。その他に、生活諸条件と体力・運動能力との相関関係等も考察している。

本学では、1983～1987年の5年間にわたって、スポーツテスト（体格・体力）の調査を実施[2]～[6]した。実施した項目は、体格診断テストの身長・体重および胸囲の3種目であり、体力診断テストの反復横跳・垂直跳・背筋力・握力・伏臥上体そらし・立位体前屈および踏台昇降運動の7種目である。この両者の診断テストを合わせると10種目である。

このデータを処理するため、使用した電子計算機は、日立製作所製のHITAC 220–IID である。この電算機と10種目に適合する統計処理のプログラムは、BASIS である。BASIS とは、Basic Statistical System の略である。BASIS は、分析の種類が豊富で、19種類の統計分析が可能で、大量のデータを取り扱うことができる。特に、回帰分析・主成分分析・因子分析・判別分析・クラスタ分析および正準相関分析などのシステム構成が可能である。

ここで、10種類のデータを表示すれば、かなりのスペースを必要とする。このスペースを小さくするため、それぞれ5種の 1 (DEGREE)、2 (DAY)、3 (MEN)、4 (GALLON) および 5 (PRODUCT 1) について置き換え、スクリーニングに関する電算処理の表示を行った。

一方、この論文に関しては、HITAC SE マニュアル・SE–1333（1991年1月）の社内報告がある[7]。その報告は「基本統計システム BASIS 適用例」である。ここでは、共同研究論文として問題点を取り上げた。

最小二乗法による線形回帰式のあてはめは、最もポピュラーな統計手法であり、パッケージの適用事例も[8]～[10]多い。しかし、入力データを吟味しないままにパッケージを適用する、いわゆる、ごみを入れてごみを出す（GIGO : Garbage In equals Garbage Out）事例が非常に多い。本来の処理目的であるトレンドや取り上げた要因効果を表わす回帰係数が正しく求められない。このため、データ特性の要約が、できないままで終了する場合がよく見受けられる。

この資料は、入手データ特性の要約に役立つ関係式（線形回帰式）をつくる

ため、実用的な事後スクリーニング法、出力結果から見た入力データ評価法を基本最小二乗法の範囲に限っている。なお、説明には、極力 BASIS マニュアルB02の例を引用し、具体的に解説したが、内容は処理システテムから独立したものである。また、リスト内容が多少異なっても、他のシステムで処理した結果についても活用することができる。

2　入力データのスクリーニング

関係式 $Y=X\beta+U$ から想像されるように、測定値の組Yは、式 $X\beta+U$ によって生成されると考えられる。そのため、Xの各列（説明変数の値の組）は、多少ともYとの比例関係が認められるものではないと、説明には役立たない。この点に一応の注意を払って、選んだXからYを近似する式：$X\beta$ を作ったとき、以下のように $Y\approx X\hat{\beta}$ の事後スクリーニングをすることができる。

入力データは *** OBSERVATIONS *** にリストされている(**Table 1**)。変数間で極端にオーダが違うと、中間データの有効けたが保たれないことがある。このような場合、パラメタ・カード ¥TRNGEN を使って、オーダの極端な不揃いを補正したり、Xを標準化する。このための変換式が成立する。[10]

$$Z_{ji}=\frac{x_{ji}-m_j}{s_j},\quad \begin{matrix} i=1,\ \cdots,\ n \\ j=2,\ \cdots,\ k \end{matrix} \quad \text{または} \quad Z=\left(I_n-\frac{1}{n}I_nI_n'\right)XS^{-1} \quad \cdots\cdots(1)$$

ここで、I_n は長さnの単位ベクトル、Sは主対角要素が第j変数 $\{x_{ji}\}$ の標準誤差 $\{s_j\}$ のk次対角行列である。なお、$\{m_j\}$ と $\{s_j\}$ は *** MEAN AND STANDARD DEVIATION *** にリストされている（**Table 2**)。このチェックは、事前に終えて置くことが望ましい。

Table 1 のDEGREE と DAY は、PRODUCT 1 と 10^2〜10^3 の開きがあり、積和行列の対角要素で 10^4〜10^6 の開きになる。掃き出し操作に移ると、内部演算（加算命令のシフト操作）で有効けたの減少が考えられる。

Table 1 では、標本の大きさ17個を処理した。その説明変数として、1 (DEGREE)、2 (DAY)、3 (MEN)、4 (GALLON) の4種類があり、被説明変数 5 (PRODUCT 1)を表示した。

次に、**Table 2** では、標本の変数別平均値と標準偏差を表示している。

Table 1 Input Data for BASIS Program
BASIS プログラムに関する入力データ

*** OBSERVATIONS ***

NO.	1	2	3	4	5
VAR.	DEGREE	DAY	MEN	GALLON	PRODUCT 1
CASE					
1	0.54500E 02	0.21000E 02	0.12900E 03	0.30670E 04	0.71070E 04
2	0.58800E 02	0.22000E 02	0.14100E 03	0.28280E 04	0.63730E 04
3	0.65200E 02	0.22000E 02	0.15300E 03	0.28910E 04	0.67960E 04
4	0.70900E 02	0.20000E 02	0.16600E 03	0.29940E 04	0.92080E 04
5	0.77400E 02	0.25000E 02	0.19300E 03	0.30820E 04	0.14792E 05
6	0.79300E 02	0.23000E 02	0.18900E 03	0.38980E 04	0.14564E 05
7	0.81000E 02	0.20000E 02	0.17500E 03	0.35020E 04	0.11964E 05
8	0.71900E 02	0.23000E 02	0.18600E 03	0.30600E 04	0.13526E 05
9	0.63900E 02	0.20000E 02	0.19000E 03	0.32110E 04	0.12656E 05
10	0.39500E 02	0.20000E 02	0.18700E 03	0.32860E 0	0.14119E 05
11	0.44500E 02	0.22000E 02	0.19500E 03	0.35420E 04	0.16691E 05
12	0.43600E 02	0.19000E 02	0.20600E 03	0.31250E 04	0.14571E 05
13	0.56000E 02	0.22000E 02	0.19800E 03	0.30220E 04	0.13619E 05
14	0.64700E 02	0.22000E 02	0.19200E 03	0.29220E 04	0.14575E 05
15	0.73000E 02	0.21000E 02	0.19100E 03	0.39500E 04	0.14556E 05
16	0.78900E 02	0.21000E 02	0.20000E 03	0.44880E 04	0.18573E 05
17	0.79400E 02	0.22000E 02	0.20000E 03	0.32950E 04	0.15618E 05

標本の大きさ：17
説明変数：1、2、3、4　被説明変数：5

Table 2 Sample Mean and Standard Deviation for BASIS Program
BASIS プログラムに関する標本の変数別平均値と標準偏差

*** MEAN AND STANDARD DEVIATION ***

No.	変数名 VARIABLE	平均値 MEAN	標準偏差 STANDARD DEVIATION
1	DEGREE	0.64853E 02	0.13510E 02
2	DAY	0.21471E 02	0.14628E 01
3	MEN	0.18182E 03	0.21995E 02
4	GALLON	0.33037E 04	0.44670E 03
5	PRODUCT 1	0.12900E 05	0.35268E 04

このように各変数名に対して、それぞれの平均値と標準偏差値が明白に区別することができる。

3　相関行列によるスクリーニング

測定値の組と説明変数との相関係数は、*** CORRELATION COEFFICIENT MATRIX *** の最後の行にリストされている (**Table 3**)。この値が小さく、一次の相関が認められないものについては、その選択を誤ったことになる。

Table 3　Sample Correlation for BASIS Program
BASIS プログラムの標本相関行列

*** CORRELATION COEFFICIENT MATRIX ***　相関行列
CORR. MTX(1, 1)

No.	1	2	3	4	5
VAR.	DEGREE	DAY	MEN	CALLON	PRODUCT 1
1	0.10000E 01				
2	0.37722E 00	0.10000E 01			
3	0.79958E-01	0.31881E-01	0.10000E 01		
4	0.35641E 00	-0.88826E-01	0.41325E 00	0.10000E 01	
5	0.12444E 00	0.10573E 00	0.91848E 00	0.63075E 04	0.10000E 01

Table 3 の DEGREE、DAY は PRODUCT 1 と一次相関があるとは認められない*。これら2変数については、代替を検討する方が得策と思われる。なお、より慣習的な統計検定法については、t 検定を参照する。

* 「本来一次相関がないのに、偶然変動だけの相関係数が作られる」
相対頻度αを見積るには、標本の大きさをnとして、

$$Z_{n-2}=\sqrt{n-2}\cdot\frac{r_n}{\sqrt{1-r_n^2}} \quad \cdots\cdots (2)$$

の値を $t(n-2,\ \alpha)$ と比べれば良い。
例えば、DEGREE については、

$$Z_{17-2}=\sqrt{17-2}\cdot\frac{0.12444}{\sqrt{1-(0.12444)^2}}\approx 0.486<t\,(17-2,\ 0.50)=0.691\cdots \quad (3)$$

$r_n=0.12444$ は、本来一次相関がなくても、偶然変動の影響だけで、1/2 以上の相対頻度で起ると考えられる。一方、GALLON については、

$$Z_{17-2}=\sqrt{17-2}\cdot\frac{0.63075}{\sqrt{1-(0.63075)^2}}\approx 3.148>t\ (17-2,\ 0.01)=2.947\cdots\cdots\cdots (4)$$

一次相関がなく、偶然変動の影響だけで、$r_n=0.63075$ が作られることは、100回に一度もない。これから DEGREE は PRODUCT 1 と一次相関はなく、GALLON は、PRODUCT 1 と一次の相関をもつと考えるのが自然である。

説明変数間の相関係数も、*** CORRELATION COEFFICIENT MATRIX *** にリストされている（**Table 3**）。特に、大きな相関係数をもつ変数の組は拮抗して、回帰係数を個別に見ると、その大きさや符号が測定値の傾向変動と合わないことがある。このような場合、測定値と説明変数の傾向変動を個々に意味付けたいならば、測定値とより高い相関係数をもつ（あるいは、より大きな t 値の絶対値をもつ(**4.3**)）説明変数で代表させるべきである。

Table 3 で DEGREE は、PRODUCT 1 と正の相関（比例）を示しているが、**Table 4** の回帰係数は負になっている。

4　諸統計量によるスクリーニング

統計量とは、標本に対応して値が決まる確立変数である。

スクリーニングに使われる統計量として、重相関係数・決定係数・F 値および t 値がリストされている。重相関係数は Y と $X\hat{\beta}$ との相関係数で、近似 $Y\approx X\hat{\beta}$ の程度を見る最も直接的な指標となる。決定係数はその一つの根拠を与える。F 値はより統計的な基準を与える。t 値は変数の説明力を個別に見るための基準を与える。

4.1　重相関係数

重相関係数（R）と決定係数（R^2）は、*** MULTIPLE REGRESSION STATISTICS *** にリストされている（**Table 4**）。R は 0.95 以上が望ましい。というのは、$R=0.95$ のとき決定係数は $R^2=(0.95)^2\approx 0.90$ であるからで

ある。これは、XとYの散布図（対応：X→Yのグラフ）の中に、Yを近似する直線 $\hat{Y}=X\hat{\beta}$ を描くとき、その直線の回りにYがランダムに散らばるためである。その変動量が、ただ単に平均値から計算した場合（$Y=X\beta+U$ という関係を想定しない場合）の1/10に減少すると解釈されるからである。

この解釈は、次の通りである。

（測定値の算術平均 $\bar{Y}$ の回りの平方和）＝（$X\hat{\beta}$ の $\bar{Y}$ の回りの平方和）
＋（$X\hat{\beta}$ 回りのYのランダムな平方和）
……(5)

が成り立つ。決定係数 R^2 は、左辺（全平方和）に対する右辺第一項の比の値である。

Table 4 Statistics Multiple Regression for BASIS Program
BASIS プログラムに関する重回帰諸統計量

*** MULTIPLE REGRESSION STATISTICS ***

No.	変数名 VARIABLE	回帰係数 REGRESSION COEFFICIENT	回帰係数の標準誤差 STD. ERROR OF REG. COE.	t 値 COMPUTED t VALUE	偏相関係数 PARTIAL CORR. COE.
1	DEGREE	-0.34658E 02	0.21492E 02	-0.16126E 01	-0.42203E 00
2	DAY	0.39686E 03	0.18615E 03	0.21320E 01	0.52414E 00
3	MEN	0.12329E 03	0.12203E 02	0.10104E 02	0.94595E 00
4	GALLON	0.29603E 01	0.66348E 00	0.44617E 01	0.78988E 00
INTERCEPT		-0.25570E 05	定数項		

COEFFICIENT OF DETERMINATION…… 0.94360E 00 決定係数
MULTIPLE CORR. COEFFICIENT…… 0.97139E 00 重相関係数
STD. ERROR OF DEPENDENT VARIABLE… 0.96713E 03 推定値の標準誤差

Table 4 では、決定係数が $R^2=0.94360$ である。これは次の直線になる。すなわち、

$$\begin{aligned}\text{PRODUCT } 1 = &-0.25570\text{E } 05 - 0.34658\text{E } 02 * \text{DEGREE}\\ &+0.39686\text{E } 03 * \text{DAY}\\ &+0.12329\text{E } 03 * \text{MEN}\\ &+0.29603\text{E } 01 * \text{GALLON} \quad \cdots\cdots\cdots(6)\end{aligned}$$

の回りを PRODUCT 1 の測定値がランダムに散らばるための変動量が、全変動量の約5.6%を占め、残りの約95%は $Y=X\beta+U$ という関係を想定することによって取り去ることができる、あるいは、説明されたと考えられる。この値は暫定モデルであるが、良い近似を示していると思われる。

4.2 分散分析のF値

F値は、*** ANALYSIS OF VARIANCE FOR THE REGRESSION *** の最後の列にリストされている (**Table 5**)。これは「測定値の組Yの変動を説明するために設けたXが、少しでも役立っている ($\beta\neq0$)」ことを主張するための根拠を与える。

$\beta=\{{}^{\beta 1}_{0}\}$、つまり $Y=X\beta+U$ のような関係がなくて、測定値がその平均値 $\bar{Y}$ の回りをランダムに変動するとしよう。そのとき、このランダムな変動の影響だけで、このような回帰係数が作られる相対頻度 α を見積るには、説明変数の個数をk、標本の大きさをnとして、このF値を $F^{k-1}_{n-k}(\alpha)$ と比べれば良い。

測定値のランダムな変動だけで $\hat{\beta}$ が作り出される相対頻度は、

$$F\text{値}\leqq F^{k-1}_{n-k}(\alpha) \quad \cdots\cdots(7)$$

のとき α より大きい。

$$F\text{値}>F^{k-1}_{n-k}(\alpha) \quad \cdots\cdots(8)$$

のとき α より小さい。

通常は、後者の場合を期待して α を、例えば0.2以下の小さな値に選ぶことができる。

Table 5 のF値50.192を $F^{k-1}_{n-k}(\alpha)$ と比較する。説明変数の個数kは、定数項も含んで $k=5$、標本の大きさは $n=17$ で、$k-1=4$、$n-k=12$ は“自由度”の列に出ている。α を例えば0.1に選ぶならば、次のようになる。

$$50.192>F^{4}_{12}(0.1)=2.48 \quad \cdots\cdots(9)$$

これから測定値のランダムな変動でけで、**Table 5** の回帰係数が作られる場合の相対頻度は、1/10よりはるかに小さいと思われる。たとえば、$\alpha=$

Table 5 Analysis of Variance Table for the Regression on BASIS Program
BASIS プログラムに関する分散分析表

*** ANALYSIS OF VARIANCE FOR THE REGRESSION *** 分散分析表

SOURCE OF VARIATION	自由度 DEGREES OF FREEDOM	平方和 SUM OF SQUARES	平均平方 MEAN SQUARE	F 値 F VALUE
回帰による DUE TO REGRESSION	4	0.18779E 09	0.46947E 08	0.50192E 02
回帰からの(残差) DEVIATION ABOUT REGRESSION	12	0.11224E 08	0.93534E 06	
TOTAL	16	0.19901E 09		

0.005 に選ぶならば、次のようになる。

$$50.192 > F^{4}_{12}(0.005) = 6.52 \quad \cdots\cdots (10)$$

となる。このことから、XはYの変動を十分説明していると思われる。

(注) F値と分散分析表の各項との関係は、次の通りである。

$$F\text{値} = \frac{\text{回帰による平均平方}}{\text{回帰からの(残差)平均平方}} = \frac{(\text{回帰による平方和})/\text{自由度}}{(\text{回帰からの(残差) 平方和})/\text{自由度}} \quad \cdots\cdots (11)$$

なお、"回帰による平方和"と"回帰からの(残土)平方和"は、それぞれ前項(5)式の右辺第一項、第二項と同じものである。

4.3 重回帰係数のt値

t値は、*** MULTIPLE REGRESSION STATISTICS *** の後ろから2列目にリストされている (**Table 6**)。これは一組の説明変数として用意したXのうち、どれが測定値の組Yの変動の説明に役立っているかを判断する基準を与える。操作的には説明力を見ようとする変数が最後に追加されたとき、ランダムな変動量がどれだけ減少するか、あるいは **Table 5** の分散分析表で、"回帰による平方和" に占める量の多少については、F表を使って調べることと同じである。

実際には、これと同等な手順としてt表を使用する。説明力を見ようとする変数の回帰係数が零であるとしよう。測定値のランダムな変動の影響だけで、

これから作り出される相対頻度αを見積るには、説明変数の個数をk、標本の大きさをnとして、この回帰係数に対応するt値をt(n−k, α)と比べれば良い。

測定値のランダムな変動だけで、**Table 6** の$\hat{\beta}$が作り出される相対頻度は、次の通りである。

$$|t\text{値}| \leqq t(n-k,\ \alpha) \quad \cdots\cdots (12)$$

のときαより大きい。

$$|t\text{値}| > t(n-k,\ \alpha) \quad \cdots\cdots (13)$$

のときαより小さい。

通常は後者の場合を期待してαを、たとえば0.2以下の小さな値に選ぶのである。

Table 6 Statistics Multiple Regression for BASIS Program
BASIS プログラムに関する重回帰諸統計量

*** MULTIPLE REGRESSION STATISTICS ***

No.	変数名 VARIABLE	回帰係数 REGRESSION COEFFICIENT	回帰係数の標準誤差 STD. ERROR OF REG. COE.	t 値 COMPUTED T VALUE	偏相関係数 PARTIAL CORR. COE.
1	DEGREE	-0.34658E 02	0.21492E 02	-0.16126E 01	-0.42203E 00
2	DAY	0.36686E 03	0.18615E 03	0.21320E 01	0.52414E 00
3	MEN	0.12329E 03	0.12203E 02	0.10104E 02	0.94595E 00
4	GALLON	0.29603E 01	0.66348E 00	0.44617E 01	0.78988E 00
INTERCEPT		-0.25570E 05	定数項		

COEFFICIENT OF DETERMINATION…… 0.94360E 00 決定係数
MULTIPLE CORR. COEFFICIENT…… 0.97139E 00 重相関係数
STD. ERROR OF DEPENDENT VARIABLE… 0.96713E 03 推定値の標準誤差

Table 6 で説明変数 DAY の |t値|=|2.1320| を t(n−k, α) と比較できる。説明変数の個数 k は、定数項も含んで k=5 とする。標本の大きさは n=17 とする。α=0.1 と α=0.05 として調べると、次の通りである。

$$2.1320 < t(12,\ 0.05) = 2.179 \quad \cdots\cdots (14)$$

$$2.1320 > t(12,\ 0.1) = 1.782 \quad \cdots\cdots (15)$$

これから説明変数 DAY の回帰係数が本来零であるのに、この測定値のランダムな変動だけで回帰係数396.86の作り出される場合は、10回中1回程度になる。したがって、説明変数DAYは、またDEGREEについては、なおYの変動を十分説明しているとは思われない。

5 おわりに

大阪経済法科大学論集〔No.41[(2)]、No.42[(3)]、No.43[(4)] および No.44[(5)]〕について、本論文との関係をまとめることができた。

それには、重相関係数と分散分析のF値である。

5.1 重相関係数

大阪経済法科大学論集〔No.41（P.83の⑤[(2)]）・No.42（P.69の⑤[(3)]）・No.43（P.87の④[(4)] および No.44（P.89の④[(5)])〕についてコメントが可能である。重相関係数は、観測値とそれに対応する回帰式の値によって計算された相関係数である。また、回帰式の当てはまりの良さを判断する直感的な尺度でもある。重相関係数は、それを2乗した決定係数が回帰式の説明力（全分散に占めるランダムでない部分の割合）を示すことからも実用上有用である。たとえば、経法論集No.42のP.69⑤について、Fの経1 R＝0.88を形式的に解釈すると、$1-R^2 \approx 1-0.8=0.2$ が説明できない。が、いわゆる観測値がランダムに回帰式の上下に変動する部分である。

5.2 分散分析のF値

大阪経済法科大学論集（No.41[(2)]・No.42[(3)]・No.43[(4)] および No.44[(5)]）について、本論文との関係をまとめるコメントが可能である。分散分析のF値は、取り上げた説明変数の有効性を判断するのに有用な尺度である。それなのに、説明変数が全く役立っていない。すなわち、回帰係数がすべて零である場合を否定できないからである。

たとえば、論集No.43のP.70②で、体重とChest（胸囲）との比例関係が全くないのに、この相関係数（0.8428）が計算されることは、1万回の観測で

も現われて来ないであろう。一方、身長と Side Step（反復横跳）とは比例関係が全くなくても、相関係数（0.0320）が計算される。このことは、3回に一度以上生じる可能性がある。このことから、体重・Chest（胸囲）のいずれかを知れば、他の値に関しては見当が付けられる。が、身長と Side Step（反復横跳）には、このような相関関係が、全くないと考えられる。

謝　辞

采山寛幸氏は、本年1月まで日立製作所関西第1システム部　技術計算グループの所属であった。そして、2月より大阪経済大学　経営情報センターの所属である。

この論文の作成に当っては、日立製作所関西第1システム部　技術計算グループと大阪経済法科大学　情報科学センターの関係各位に対して厚くお礼を申し上げます。

参 考 文 献

（1）文部省体育局「昭和59～平成2年度体力・運動能力調査報告書」文部省体育局（1984～1990）

（2）勝　英雄、沢　勲「1987年度大阪経済法科大学学生のスポーツテスト・データ（体格・体力）に関する統計的分析」大阪経済法科大学論集 **41** P.21～84（1990）

（3）沢　勲、大森敏行「1986年度大阪経済法科大学学生のスポーツテスト・データ（体格・体力）に関する統計的分析」大阪経済法科大学論集 **42** P.5～70（1990）

（4）沢　勲、森下泰行「1985年度大阪経済法科大学学生のスポーツテスト・データ（体格・体力）に関する統計的分析」大阪経済法科大学論集 **43** P.29～89（1990）

（5）沢　勲、高垣英夫「1984年度大阪経済法科大学学生のスポーツテスト・データ（体格・体力）に関する統計的分析」大阪経済法科大学論集 **44** P.35～91（1991）

（6）沢　勲、中澄孝司、高垣英夫「大阪経済法科大学学生（1983～1987年）の立位体前屈と踏台昇降運動の体力統計」大阪経済法科大学総合科学研究所年報 **10** P.23～58（1991）

（7）（マニュアル）「プログラムプロダクト VOS 0/VOS1-S/VOS1 基本統計システム BASIS（資料番号 340-7-031-20）」株式会社日立製作所（昭和55年）

（8）J. Johnston「Econometric Methods」McGraw-Hill（1963）
（竹内　啓訳「計量経済学の方法」東洋経済新報社　昭和49年）

（9）森口繁一（編）「日科技連数値表（A）」日本科学技術連盟（1969）

（10）芝祐順「行動科学における相関分析法」東京大学出版会（1975）