



慶應義塾大学ビジネス・スクール

MEノート（４）

“ REG ” について

以下はBASICのシステムに入ってから解説である。MS-DOSシステムの場合は先ず附録（８ページ以降）を参照のこと。

1. このプログラムは回帰分析を行うプログラムである。通常のパラメーターの出力の他に、簡単な変数変換（変数の除去を含む）と予測、および簡単な残差のチェックを行うことができる。

2. 実行に当っては先ずMEノート(3)の1～3をよく読んでデータファイル“data0”を作成する。その手順は先ず、

```
load “D=GEN”®
```

としてライン番号1～4999にデータ文で自分の分析したいデータを入力する。それが終わったら自由に名称をつけて、例えば“ABC”で

```
SAVE “ABC”®
```

として別に確保する。

次いで

```
RUN®
```

とすると“data0”が作成される。

次に

```
load “REG”®
```

```
RUN®
```

とするとプログラムの実行が開始される。

以下に実行の際の画面での応答の仕方と出力について、実行の際の順序に従って簡単に説明する。

3.1) 入力データの確認

実行開始と同時に、入力データの確認が行われる。画面に2行の数字（変数の数が多いときには何行かにわたる）が打出される。最初の行はデータ番号1，第2行は最後のデー

タ番号に対応しており，左から

被説明変数，第一変数，……，第K変数

の順序になっている。自分の意図したデータが正しく入力されているかどうかをこゝで確認する。もし正しく入力されていれば，

OUTPUT ALL DATA ?

に対して 0Ⓡ と応答する。若し正しく入力されていなければ，こゝで 1Ⓡ を入力してデータの全リストを取りそれを手がかりにしてもう一度 "D=GEN" のデータを修正する。

ロ) 簡単な変数の変換

次いで画面は

ENTER TRANSFORMATION ?

となるので入力したデータのまゝで回帰分析を行うときは 0Ⓡ と応答すると，下記へ以下に進む。1Ⓡ と応答すると画面はライン番号 7000 台の変換のサブルーチンに入って，変換コードを表示する。

TRANSFORMATION CODES FOR Y&X "

' 1 ' FOR AS IS "

' 2 ' FOR STANDARDIZATION "

' 3 ' FOR V=LOG (V) "

' 4 ' FOR V=V^2 "

' 5 ' FOR V=1./V "

' 6 ' FOR V=0 "

' 7 ' FOR V(I) = (V(I) - V(I-1)) / V(I-1)

' 1 ' は何も変換を施さないことを意味する。

' 2 ' は変数を標準化する。すなわちその変数について

$$\frac{\text{変数の値} - \text{その変数の平均値}}{\text{その変数の標準偏差}}$$

その変数の標準偏差
の形の標準化を行う。その際使用した平均値と標準偏差がプリンターに出力される。

' 3 ' は対数変換を施す。対数の底は 10 である。

' 4 ' はその変数の値を 2 乗する。

'5' は逆数をとる。

'6' はその変数を説明変数から除外する。

'7' は時系列でデータが入力されているときに成長率を与える。この変数の際にはデータ数が1だけ自動的に減少する。

5

続いて画面は

TYPE IN THE DESIRED CODES FOR Y AND XS

Y = ?

となるので先ずYについての変換コードとして、例えば

1 Ⓡ

10

と入力する。続いて画面は、

X (1) = ?

となるのでこれに対しても例えば 1 Ⓡ と入力する。次の

X (2) = ?

に対しても同様に例えば 6 Ⓡ と入力する。以下同様に全ての変数について変換コードを指定する。

15

上述の例では本来は

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

であったモデルが

$$Y = \beta_0 + \beta_1 X_i + u$$

20

となる。また

Y = ? 1 Ⓡ

X (1) = ? 3 Ⓡ

X (2) = ? 3 Ⓡ

と入力した場合は同様に

25

$$Y = \beta_0 + \beta_1 \log_{10} X_1 + \beta_2 \log_{10} X_2 + u$$

というモデルに関して回帰分析を行うことになる。

変換が終ると画面に

FURTHER TRANSFORMATION ?

30

が出る。これは例えば一度標準化したものゝ対数をとりたいといった場合に 1 Ⓡ とする。これによって様々な形の変換が可能となる。これに対して 0 Ⓡ とすれば、画面は

変換されたデータを確認のために出力する。この場合、以前と同様に最初と最後のデータの組だけが画面に現れるので、予定通りに変換されているかどうか、又変数の位置も予定通りであるかどうかを確認する。

次いで画面は

SAVE NEW DATA ?

となる。修正したデータを保存したいときは、1 とする。このとき作成されるファイル名は "data1" となっている。不要なときは 0 と入力すると次のステップに進む。

ハ) 各種パラメーターの計算

計算が開始されてしばらくすると、順調なら

WORKING WELL SOMEHOW

というメッセージが現れる。もしこれが現れないときは入力データがおかしいのもう一度 "D = GEN" で調べて見るのがよい。(例えば殆んど相関係数が1である様な二つの変数を説明変数にしているなど……)。

計算が終了すると次の様なアウトプットを得る。

* PARAMETER INFORMATION

VARIABLE	ESTIMATES	STD. DEV.	T-VALUE
0	1.8741	1.0606	1.7670
1	0.5524	0.4238	1.3037
2	0.8182	0.3706	2.2079

COEFFICIENT OF DETERMINATION = .848951
(DURBIN-WATSON STATISTIC = 2.6015)

VARIABLE は変数名である。0は定数項を示す。

ESTIMATES は回帰係数の推定値である。上から順に

$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ の値である。

STD. DEV は $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ の標準偏差の推定値である。

T-VALUE は $(\hat{\beta}_i - 0) / (\hat{\beta}_i \text{の標準偏差})$ の値である。これは $\beta_i = 0$ という仮説

を検定するとき利用する。データの数が少ないときは自由度を（データ数－説明変数の数－1）としてt分布の数表を用いる。データ数が大きいとき（自由度20以上）のときは正規分布の数表を用いる。

COEFFICIENT OF DETERMINATION これは所謂決定係数であり，被説明変数の変動のうち説明変数の変動によってどれだけの比率が説明されるかを示している。勿論この値が1に近いほど説明される程度は大きい。しかしそれが意味のある説明かどうかは別問題である。

(DURBIN-WATSON STATISTIC) これはデータが時系列であるときに残差項の間に時系列的に関連性があるかどうかを見るためのものであり，0～4の間の値をとる。両端の値に近いときは攪乱が互いに独立ではない可能性が大きいので注意を要する。データが時系列でないときは無視すること。

二) 予 測

次に画面に

ENTER PREDICTION ?

というメッセージが現れる。予測するときは 1Ⓡ，そうでないときは 0Ⓡ を入力する。このときホ)へ進む。

1Ⓡ を入力すると画面は

PREDICTION TRIAL NO.1

TYPE IN X-VALUES

X (1) = ?

となって停止する。こゝでX (1)の条件値を，例えば 3.5Ⓡ と入力する。以下同様にしてすべての説明変数について条件値を入力する。入力する値は変換前の値でよい。

全ての条件値を入力し終ると入力データが

X-VALUES

としてプリンターに記録され，ついで計算結果

PREDICTED Y-VALUE =.....

STD. DEV. OF PREDICTOR =.....

が画面とプリンターの双方に現れ，

ANOTHER TRIAL ?

というメッセージとなる。条件値を変えて再度予測したいときには 1Ⓡ，そうでないときは 0Ⓡ と応答する。

ホ) 次のステップでは残差項に関する検討を行うことができる。

先ず

RESIDUAL EXAMINATION ?

に対して 1 又は 0 で応答する。0 のときは全プロセスが終了する。

1 と応答した場合には

OUTPUT OF RESIDUALS ?

が現れる。1 のときは次の様なアウトプットを得る。

* OUTPUT OF RESIDUALS

E (J)	Y (J)	Y-HAT	A/F RATIO
-0.7972	3.0000	3.7972	0.7901
0.7552	4.0000	3.2448	1.2328
-0.4336	5.0000	5.4336	0.9202
0.2797	6.0000	5.7203	1.0489
0.1958	7.0000	6.8042	1.0288

こゝで E(J) は J 番目のデータに対応する残差項、Y(J) は被説明変数の値、Y-HAT はデータの説明変数の値を条件値とするときの推定値、A/F は Y(J)/Y-HAT で実績値と予測置の比である。この A/F 比の分布と以前に得たパラメーターの推定値から別の方法で予測を行うこともできるであろう。

へ) 残差項の性質を検討するためのグラフ

次に画面の表示は

SCATTER DIAGRAM ?

となる。こゝでは "D=GEN" で用いたと同じプログラムを用いて縦軸を残差項の値として散布図を画く。こゝで 1 と応答すると、

TIME SERIES ?

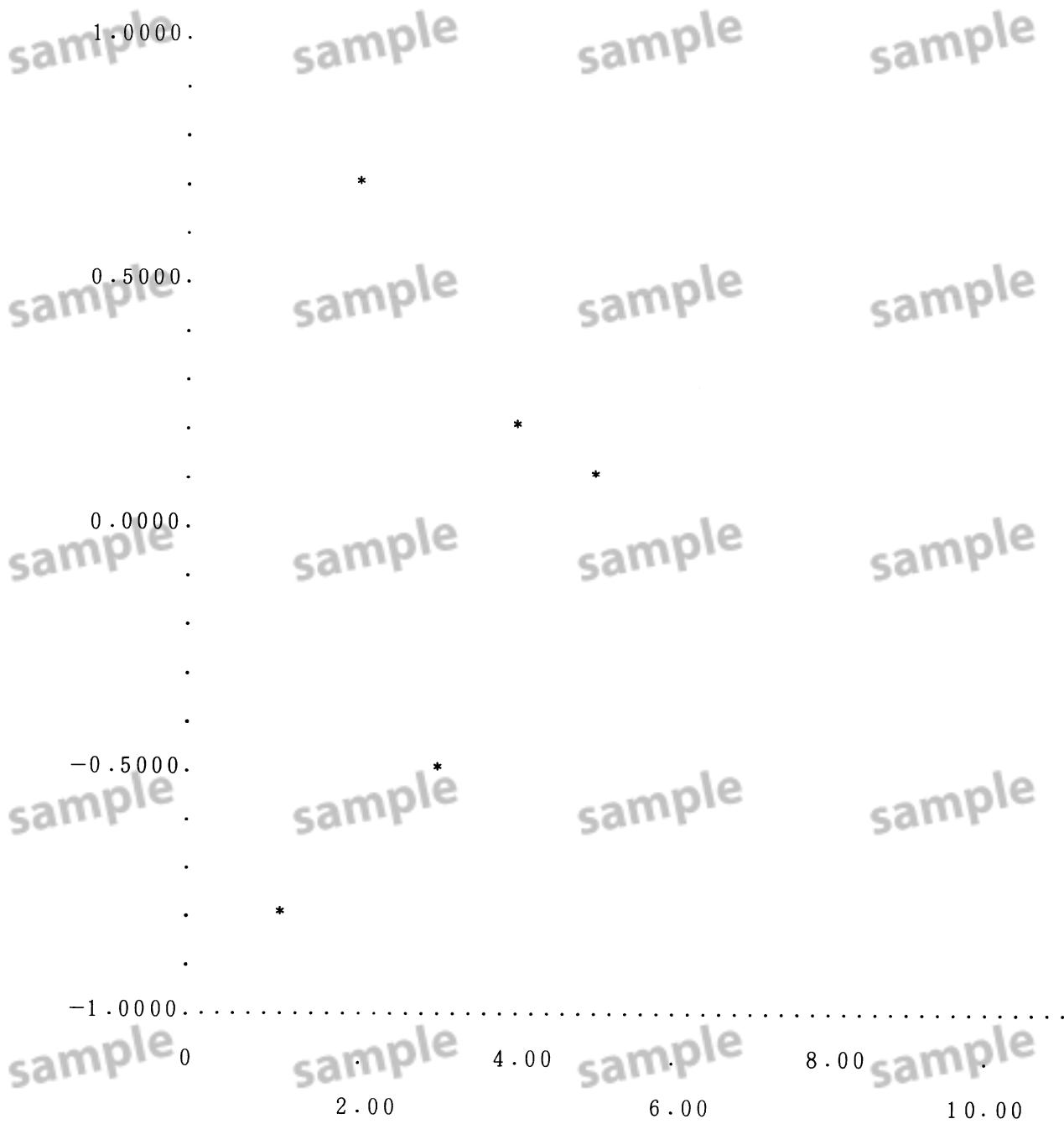
となる。これは横軸を時間軸にするかどうかを訊ねている。1 とすると直ちに、

TYPE IN X-MAX, X-MIN, E-MAX, E-MIN ?

が表示される。ME ノート(3), 4.ニ) の要領で横軸の両端の値と縦軸の両端の値 (今は E (J) の値) を入力する。例えば、

0, 10, -1, 1

を入力すると次に見られる様なグラフが得られる。



附録

MS-DOS版「DGEN」及び、「REG」の利用法

ディスク内にN88BASIC.EXEとN88BASIC.LIBの2つのファイルがあるのを確認してから、

BASIC ®

と入力すると、画面に通常の BASIC と同様のメッセージが現われる。後の操作は、通常の場合と、殆んど、全く同一である。違いは、ファイル名が「=」が落ちていること、システムのメモリーが少し小さくなっていること位である。

例えば、「DGEN」を使用するには、

load "DGEN" ®

run ®

とすれば良い。MS-DOS版ではファイル名の大文字小文字の区別をしていないから、上述のコマンドは、load "dgen" ® でも、同じ動きをする。

「DGEN」でファイル作成が終わったら、そのまま「REG」の使用に進んでも良いが、

SYSTEM ®

としてBASICを抜けてMS-DOSのシステムへ戻り、次いで、単に、REGと入力しても良い。この最後の方法は、次の様に、各種のスプレッド・シートで作成されたデータを利用する際に、合わせて用いられる。

VP-plannerのデータによる「REG」の利用

1. 数値データのためのファイルの場合

数値1にVP-plannerのシステム・ディスクを挿入し、装置2にMS-DOSのファイルを入れ、スイッチをONにして、システムを立ち上げる。

VP-plannerのワークシートに入ったところで、各々の場所を連続して使用して、例えば次の様に、「D=GEN」の場合と同一形式のファイルを作成する。

	A	B	C
1	4	3	
2	12	8	17
3	7	11	9
4	4	9	6
5	8	6	11
.			
.			
.			

データの入力が終わったら、まず、/FSで、データの性格を示すファイル名を使用して、ワークシートを保存する。

次に、/FXで、ファイル名を「data0」とし、S「標準」を選択し、「範囲」をA1...C5として部分保管する。これで「REGVP」への入力用のファイルが作成された。

装置1のファイルを「REGVP」の入っているファイルと交換し、リセット・ボタンを押して、もう一度システムを立ち上げる。

次に、

copy b:data0.prn @

として、B:にあるデータ・ファイルをA:に転送する。

後は、

regvp @

とすれば、「REG」のマニュアルと全く同じ動作をする筈である。

2. ワークシートの中に変数名等が記入されている場合

典型的にはワークシートの内容は次の様なものが多いであろう。

	A	B	C	D	...
1	地域別製品売上データ				
2		4	3		
3	地域番号	A製品	B製品	C製品	
4	1	12	8	17	
5	2	7	11	9	
6	3	4	9	6	
7	4	8	6	11	
.	.				
.	.				
.	.				

こゝで、B 2 のデータ数に対応するデータ 4 と、C 3 の変数の数に対応するデータ 3 が無い場合には、新たにこの場所に行を追加して、この形にしなくてはならないであろう。この形のワークシートが完成されたならば、上記 1. の手続きで、部分保管の時に「範囲」B 2.. D 7 とし、「REGVP」ではなく、「REGVP1」を使用することにすれば、後は全く同一である。

Multiplan, Lotus 1-2-3, Supercalc
等 (バージョンを問わない), のワークシート・データの使用

それぞれのワーク・シートで, データの種類に応じて, 1.域いは, 2.での範囲を指定して,
P「印刷」, F「ファイル」を選択し, ファイル名を, 例えば, 「sales.prn」として, 出力
してファイルを作成する。

このファイルは, 拡張子が .prn となっているので, VP-planner のワーク・シートに,
/FI「ファイル, 読み込み」で, 読み込むことができる。

後は, 上記 1., 或いは, 2.の手続きに従えば良い。

不 許 複 製

慶應義塾大学ビジネス・スクール

Contents Works Inc.