

# SAS Technical News

Summer 2002

*For Higher  
Customer Satisfaction,  
We Bridge  
the SAS System  
Between  
Customer's World.*

## CONTENTS

- 1 特集 変数のクラスタリング PROC VARCLUS再発見 (再掲)
- 8 Q&A
- 13 SASトレーニングのお知らせ
- 14 SASプログラマーズ・ワークショップ開催のお知らせ
- 14 新刊マニュアルのご紹介
- 15 Web Based Trainingのご紹介
- 16 最新リリース情報

## 特集

### 変数のクラスタリング PROC VARCLUS再発見 (再掲)

#### 1. 2002年TechNews版結言

たとえば質問紙調査の構成的妥当性を調べるため、因子分析がよく適用される。主な目的は変数群をクラスターに分けることである。因子分析はなかなか奥が深い方法論なので、応用上の幅は広がるが、適用上克服していかねばならないさまざまな困難が存在する。単に変数群をクラスター分けするだけなら、それ専用の方法を適用した方がより簡潔でまた明快であるといえないだろうか。PROC VARCLUSは、まさにそのような要求に応えるプロシジャであり、多変量型データの変数の方、すなわち相関行列または分散共分散行列を直接クラスタリングするものである。実用上有益な方法論だと思われるが、これまで応用例が発表されていなかった。このたびAsukai et al.(2002)によりPTSD尺度構成にこのVARCLUSを利用した例が精神医学関連の国際雑誌に掲載された。今回の特集では、このVARCLUSを使った論文発表を記念して、そのアルゴリズムと出力の解釈・指定方法について解説した岸本(1996)の論文を再掲する。1996年発表の論文なので、SASシステムのリリース関連の情報が当時のままになっている点にご留意いただきたい。文中のプログラムの実行及び分析処理結果についてはWindows版SASシステムバージョン6のものであるが、同じプログラムをWindows版SASシステムバージョン8でもエラー無しで実行できる。

#### 2. はじめに

「分類」は研究の初期において非常に重要な手段である。SASのPROC CLUSTERに収載されているようなクラスター分析の諸手法は、個体間の距離行列から分類をスタートさせる。多変量型のデータにおいて、個体間の距離は通常ユークリッド距離(の2乗)で定義される。すなわち、通常のクラスター分析の方法は多変量型データのオブザベーションを分類するために使われることが多い。多変量型のデータが与えられて、変数の方を分類したいときもある。このときは、実際には相関行列あるいは分散共分散行列から変数をクラスタリングするというのが課題となる。変数のクラスタリングについて、伝統的には次のような方法が採られている。

- ・ 相関係数を操作して距離とみなせる数値を作り、そこから通常のクラスター分析を行なう。たとえば、相関が1のとき距離がゼロで、-1のとき最大の距離になるようにするには、 $d=1-r$ と変換することが考えられる。ただし、このように計算した距離は、いわゆる距離の公理をみたさない。
- ・ 因子分析を行い、主観的にクラスタリングを行なう。この方法はファジイなクラスタリングなので現実的な分類がやや難しい。方法論的にもやっかいな問題がある。また、直交解は現実のデータに合わないことが多い。

PROC VARCLUSは、SASに以前から存在する変数のクラスタリングのためのプロシジャである。このプロシジャは、相関行列または分散共分散行列を入力データとして、その変数としての特質を活かしながら直接的なクラスタリングを行なう。方法論的には、分割型(最初はデータ全体が1つのクラスターで、だんだん分割していく方法)に属する。必ずしも階層型ではないが、階層型の制約を課すことも可能である。明確に分離した分類結果が得られるため、結果の解釈は容易である。このように良い特性を持つアルゴリズムであるが、多変量解析の教科書や論文に記載がないため、広く使われているとはいえない。本稿では、この隠れた才能を発掘して紹介する。

### 3. アルゴリズムの紹介

#### 3.1 例題データ

アルゴリズムを説明するための例題として、次のような5変数の人工的データを用いる。

```
data VCTEST;
  input X1-X5;
  cards;
10 9 9 8 9
8 10 4 6 1
3 0 8 9 3
1 2 3 2 6
;
```

分析は相関行列からスタートすることにする。PROC CORRで相関行列を求めて表示する。

```
proc corr data=VCTEST nosimple noprob;
run;
```

Pearson Correlation Coefficients / N = 4					
	X1	X2	X3	X4	X5
X1	1.00000	0.89765	0.45796	0.49955	0.19260
X2	0.89765	1.00000	0.02268	0.10246	0.06201
X3	0.45796	0.02268	1.00000	0.87782	0.42056
X4	0.49955	0.10246	0.87782	1.00000	-0.05384
X5	0.19260	0.06201	0.42056	-0.05384	1.00000

変数X1-X2間、およびX3-X4間の相関がやや大きい。全体的に正の相関をとることが多い一群のデータである。

#### 3.2 分割するための基準

まず、データ全体の変数群を1つのクラスターであると考えよう。ここから、さらにクラスターに分割する価値があるか否かを判断するため、次の基準を導入する。

第2主成分( PRIN2 )の固有値( Eigenvalue )が、たとえば1( 相関行列から分析を始めた場合 )以上のとき

クラスターを代表する成分としては、主成分よりも重心成分( すべての変数についての重みが等しい平均ベクトル )の方が適切かもしれない。重心成分を使うときは固有値は計算しないので、

重心成分で分散を説明する寄与率が、たとえば75%以下のとき

という基準を用いることも考えられる。もちろん第1主成分の寄与率を用いてもよいが、主成分を計算するのであれば第2固有値の基準を用いた方がよいであろう。PROC VARCLUSでは、どちらの基準も利用できるが、デフォルトは前者である。今のデータを第2固有値の基準に照らして、分割する価値があるかどうか判定してみよう。

```
proc princomp data=VCTEST;
run;
```

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	2.53417	1.11086	0.506833	0.50683
PRIN2	1.42331	0.38078	0.284662	0.79149
PRIN3	1.04253	1.04253	0.208505	1.00000
PRIN4	0.00000	0.00000	0.000000	1.00000
PRIN5	0.00000	.	0.000000	1.00000

PROC PRINCOMPで主成分分析を行なった結果、第2主成分の固有値は1.42であり、クラスターに分割するべきであると判定された。第1主成分の固有値は2.53、寄与率は50.7%であり、このことから変数全体を1クラスターとみなすことはできないことが示唆される。

#### 3.3 第1次割りつけ

全変数群を2つに分割することを考える。あらかじめ分割すべきクラスター数がわかっていれば直接その数に分割することもできるが、クラスター数が不明のときには各クラスターが前節の基準を満たすように1つずつクラスター数を増加させていく方式がとられる。変数の分割方法は、k-menasクラスタリング法と類似している方法が用いられる。まず、分割のシードとなる初期クラスター成分を設定する。初期クラスター成分は、原理的にはなんでもよく、たとえば元データの変数群のうち最初の変数と次の変数とを選んでよい。しかし、これから作成する変数クラスターの中心を貫通するような良い初期クラスター成分を選んでおくと収束が速く、また局所最適に陥る可能性も少ない。PROC VARCLUSで採用されている初期クラスター成分の生成法では、2番目までの主成分にHarris-Kaiser( 1964 )のオーソプリク回転という斜交回転法を施す。オーソプリク回転は、PROC FACTORにmethod=HKを指定しても実行させることができる。

```
proc factor data=VCTEST nfactors=2 rotate=HK;
run;
```

Rotation Method: Harris-Kaiser		
Factor Structure (Correlations)		
	FACTOR1	FACTOR2
X1	0.51532	0.96585
X2	0.08628	0.98115
X3	0.99706	0.21332
X4	0.90056	0.28147
X5	0.38416	0.11490

Factor Structure(因子構造)とは、斜交回転した因子(この場合初期クラスター成分)と各変数との相関係数である。この値を基準として、各変数がどちらのクラスターに所属するかを割りつけていく。因子構造は負の値をとることもあるので、近さを表すには2乗するとよいであろう。今の例では、変数X3-X5までが第1クラスターに、X1とX2とが第2クラスターに割りつけられる。

### 3.4 第2次割りつけ

初期クラスター成分への割りつけをさらに改良するため、クラスター成分を計算しなおして第2次割りつけを行なう。新しいクラスター成分には、各クラスターに属する変数群の第1主成分(デフォルト)か、あるいは重心成分が用いられる。ここでは一貫して第1主成分を用いる。新しいクラスター成分への相関係数(の2乗)を基準に、各変数がどちらのクラスターに属するかを判定しなおす。

```
proc princomp data=VCTEST n=1
  out=CLUSTER1(rename=(prin1=CLUS1));
  var X3 X4 X5;
run;

proc princomp data=CLUSTER1 n=1
  out=CLUSTER2(rename=(prin1=CLUS2));
  var X1 X2;
run;

proc corr data=CLUSTER2 nosimple noprob;
  var X1-X5;
  with CLUS1 CLUS2;
run;
```

Pearson Correlation Coefficients / N = 4					
	X1	X2	X3	X4	X5
CLUS1	0.50193	0.07103	0.99818	0.89721	0.38972
CLUS2	0.97408	0.97408	0.24672	0.30902	0.13069

第1クラスターに所属する変数はX3-X5で、第2クラスターに所属する変数はX1とX2と判定される。この結果は初期クラスター成分を用いた分類結果と変わらない。ゆえに、2つのクラスターへの分割過程は収束したとみなされる。以上のように、オブザベーションの分類におけるk-means法に対比される変数の分類法を、NCS(Nearest Component Sorting)フェーズと呼ぶ。この後、さらに探索フェーズと呼ばれる過程を実行することができる。これは、各変数のクラスターの割りつけを変更して、分散説明率が上昇するかどうかをテストするものである。分散説明率が上昇したときは、その都度クラスター成分を再計算する。主成分をクラスター成分にした場合、NCSフェーズを反復すると分散説明率は単調に増加(正確には非減少)し、しかも高速に収束する。オーソブリク回転による初期クラスター成分を用いた場合、NCSフェーズの後に探索フェーズを追加しても分類結果がさらに改善されることはめったにない。したがって、主成分を使うときのデフォルトでは、探索フェーズは実行しない。一方、重心成分をクラスター成分としたときには、NCSフェーズを反復しても分散説明率は必ずしも増加しない。したがって、重心成分を使うときのデフォルトでは、NCSフェーズを1回だけ実行し、その後探索フェーズを10回反復する。

### 3.5 さらに細かい分割

第1段階で2分割したクラスターそれぞれについて主成分分析を行なった結果を示す。

```
proc princomp data=VCTEST; /* クラスター1 */
  var X3 X4 X5;
run;
```

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	1.95322	0.91132	0.651075	0.65107
PRIN2	1.04191	1.03704	0.347302	0.99838
PRIN3	0.00487	.	0.001623	1.00000

```
proc princomp data=VCTEST; /* クラスター2 */
  var X1 X2;
run;
```

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	1.89765	1.79529	0.948823	0.94882
PRIN2	0.10235	.	0.051177	1.00000

第1固有値の値を合計すると、 $1.95+1.90=3.85$ であり、変数全体で主成分分析した場合の固有値2.53よりも分散説明率がかなり向上していることがわかる。さらに細かくクラスターに分割する価値があるか否かを判定するために、各変数クラスターについてクラスター分離の基準(第2固有値が1以上)を適用してみる。すると、クラスター1の第2固有値は1.04であり、さらに分割できることが示唆される。クラスター1の変数群について、これまでに紹介したようなオーソブリク回転による初期クラスター成分設定NCSフェーズによる変数割りつけで変数群を2分割する過程を再帰的に実行する。変数群をクラスターに割りつけるとき、デフォルトでは全変数があらためて各クラスター所属への候補となる。特に指定をすれば、階層型の制約を与えることができる。すなわち、前の段階で分割されたクラスターをさらに分割するときには、前の段階のクラスターを構成する変数のみが次の分割の所属候補となる。クラスター分割の過程をデンドログラムで表現したいとき、または計算時間を節約したいときは、階層型の制約をつけるとよい。変数を3分割したときのクラスター成分と各変数との相関について結果だけ示すと、次のようになる。

Pearson Correlation Coefficients / N = 4					
	X1	X2	X3	X4	X5
CLUS1	0.49408	0.06458	0.96897	0.96897	0.18923
CLUS2	0.97408	0.97408	0.24672	0.30902	0.13069
CLUS3	0.19260	0.06201	0.42056	-0.05384	1.00000

クラスター1には変数X3,X4が、クラスター2には変数X1,X2が、クラスター3には変数X5が所属する。この状態で分散説明率は4.78まで増加する。これ以上の分割はできないので、3つの変数クラスターを得た段階で分割過程を終了する。

## 4. PROC VARCLUSの出力と解釈

### 4.1 PROC VARCLUSのプログラム

例題データを対象にPROC VARCLUSを実行して、第2章で説明したアルゴリズムがどのように表現されているか、また出力をどのように解釈したらよいかを調べてみよう。PROC VARCLUSのプログラムは、次のようにオプション指定を最小限にしたものである。

```
proc varclus data=VCTEST;
run;
```

### 4.2 データ全体の情報

```
Oblique Principal Component Cluster Analysis

4 Observations      PROPORTION =      0
5 Variables         MAXEIGEN  =      1

Cluster summary for 1 cluster(s)

Cluster      Members  Variation  Proportion  Second
-----
Cluster 1    5        5.00000   2.53417    0.5068     1.4233

Total variation explained = 2.534166 Proportion = 0.5068
Cluster 1 will be split.
```

まず、斜交主成分クラスター分析というタイトルが表示される。これから、クラスター成分として重心成分ではなく主成分を用いたことがわかる。次に、データのサイズとクラスター分割の基準が表示される。PROPORTIONの値は、クラスター成分の分散説明率の基準をあらわす。たとえばクラスター成分がデータ分散の80%を説明したら分割を止めるときには、PROPORTION=0.8とオプション指定する。主成分を用いたときのデフォルト設定はゼロなので、分散説明率の基準は用いられない。その下のMAXEIGEN=の値は、分割を止めるときの第2固有値の基準を示す。主成分を用いたときのデフォルトは1なので、第2固有値が1以上ならさらに分割することになる。次に、データ全体をこれから分割するための基礎資料が表示される。現在、全変数がクラスター1に所属するので、クラスター1のメンバーは5である。分析を相関行列からスタートさせるので、各変数の分散は1に基準化され、クラスター全体がもっている変動の大きさは5である。第1主成分の固有値が2.53なので、第1クラスター成分は全体の分散の51%を説明した。また、第2固有値の大きさが1.42なので、変数全体のクラスターは分割される。

### 4.3 2分割したクラスターの情報

```
Oblique Principal Component Cluster Analysis

Cluster summary for 2 cluster(s)

Cluster      Members  Variation  Proportion  Second
-----
Cluster 1    3        3.00000   1.95322    0.6511     1.0419
Cluster 2    2        2.00000   1.89765    0.9488     0.1024

Total variation explained = 3.850869 Proportion = 0.7702

R-squared with
-----
Variable  Cluster  Closest  1-R**2
-----
Cluster 1-----
X3         0.9964   0.0609   0.0039
X4         0.8050   0.0955   0.2156
X5         0.1519   0.0171   0.8629
Cluster 2-----
X1         0.9488   0.2519   0.0684
X2         0.9488   0.0050   0.0514
```

2番目のページには、変数全体を2つのクラスターに分割したときの資料が示されている。第1クラスターの第2固有値は1.04であるから、このクラスターはさらに分割される。2つのクラスターの固有値を合計すると3.85であり、データ全体の分散の77%を説明している。その下方には、変数全体を2分割したときの各変数についての資料が表示される。X3-X5までが第1クラスターに、X1とX2とが第2クラスターに属すると判定された。各変数とクラスター成分との近さの指標として、相関係数の2乗が用いられる。その変数自身が属しているクラスターと相関が高く、近接している別のクラスターとは相関が低いのが、クラスターメンバーとして良い変数である。参考のため、 $(1-R^2_{own}) / (1-R^2_{next})$ の比が表示されている。この値が大きいと、クラスターメンバーとしては適切でないことがわかる。たとえば、変数X5はクラスター1のメンバーとしてはふさわしくないことがわかる。

### 4.4 2分割したクラスター成分の詳しい情報

```
Standardized Scoring Coefficients

Cluster      1          2
-----
X1           0.000000    0.513307
X2           0.000000    0.513307
X3           0.511042    0.000000
X4           0.459347    0.000000
X5           0.199526    0.000000
```

Cluster Structure

Cluster	1	2
X1	0.501930	0.974075
X2	0.071031	0.974075
X3	0.998179	0.246716
X4	0.897208	0.309017
X5	0.389720	0.130691

Inter-Cluster Correlations

Cluster	1	2
1	1.00000	0.29410
2	0.29410	1.00000

Cluster 1 will be split.

3ページ目には、クラスター成分について詳しい情報が表示される。クラスター成分は、たまたま得られた各クラスターの変数群について別々に主成分分析を行ったものであるから、最下部に示されているように、クラスター成分間には自然な相関関係が存在する。そのため、各成分の性質は(因子分析の斜交回転のように)複雑なものになる。Standardized Scoring Coefficientsというのは、元の変数群からクラスター成分を合成変数として計算するときの係数である。この値は、PROC FACTORでSCOREオプションを設定すると再現できる。Cluster Structureというのは、クラスター成分と各変数との相関係数である。この値を2乗すると、前ページのR<sup>2</sup>の値が得られる。変数の分類のみに興味があり、クラスター成分の性質についてあまり興味がないなら、このページの出力は軽視してさしつかえない。これらの情報は、PROC VARCLUSステートメントにSHORTオプションをつけると表示されなくなる。

4.5 3分割したクラスターの情報

Cluster summary for 3 cluster(s)

Cluster	Members	Cluster Variation	Variation Explained	Proportion Explained	Second Eigenvalue
1	2	2.00000	1.87782	0.9389	0.1222
2	2	2.00000	1.89765	0.9488	0.1024
3	1	1.00000	1.00000	1.0000	.

Total variation explained = 4.775466 Proportion = 0.9551

R-squared with

Cluster	Variable	R-squared with		
		Own	Next	1-R**2
Cluster 1	X3	0.9389	0.1769	0.0742
	X4	0.9389	0.0955	0.0675
	X5	0.9389	0.0955	0.0675
Cluster 2	X1	0.9488	0.2441	0.0677
	X2	0.9488	0.0042	0.0514
Cluster 3	X5	1.0000	0.0358	0.0000

Standardized Scoring Coefficients

Cluster	1	2	3
X1	0.00000	0.51331	0.00000
X2	0.00000	0.51331	0.00000
X3	0.51601	0.00000	0.00000
X4	0.51601	0.00000	0.00000
X5	0.00000	0.00000	1.00000

Cluster Structure

Cluster	1	2	3
X1	0.49408	0.97408	0.19260
X2	0.06458	0.97408	0.06201
X3	0.96897	0.24672	0.42056
X4	0.96897	0.30902	-0.05384
X5	0.18923	0.13069	1.00000

Inter-Cluster Correlations

Cluster	1	2	3
1	1.00000	0.28676	0.18923
2	0.28676	1.00000	0.13069
3	0.18923	0.13069	1.00000

No cluster meets the criterion for splitting.

3分割したときの情報も2分割のときと同様に現れる。3つのクラスターのいずれもさらに分割する基準を満たしていないので、分割過程はここで終了する。

4.6 クラスタリング履歴

Oblique Principal Component Cluster Analysis

Number of Clusters	Total Variation Explained by Clusters	Proportion of Variation Explained by Clusters	Minimum Proportion Explained by a Cluster
1	2.534166	0.5068	0.5068
2	3.850869	0.7702	0.6511
3	4.775466	0.9551	0.9389

Number of Clusters	Maximum Second Eigenvalue in a Cluster	Minimum R-squared for a Variable	Maximum 1-R**2 Ratio for a Variable
1	1.423308	0.1098	.
2	1.041906	0.1519	0.8629
3	0.122180	0.9389	0.0742

最後に、クラスター分割を進めていった履歴が表示される。分割が進むにつれて、各クラスターの固有値の合計、つまり説明される分散が増加していく。データ全体の分散は同じであるから、各クラスターで説明される分散の割合も増加していく。この値で全体的なクラスタリングの良さが評価できる。うまくクラスタリングできたなら、分散説明率はかなり高く(たとえば80%以上に)なっているはずである。次に、各クラスターの中で最もまとまりが悪いクラスターについて、分散説明率・第2固有値・そのクラスターに

属している各変数との $R^2$ のうち最小のもの $\cdot(1-R_{own}^2)/(1-R_{next}^2)$ の値の中で最大のものの4つの統計量が表示される。これらの値が悪ければ、さらにクラスター分割を進めることが示唆される。

## 5. PROC VARCLUSの指定法

### 5.1 重心成分を使ったクラスタリング

PROC VARCLUSの指定方法には、さまざまなバリエーションがある。たとえば、デフォルトではクラスター成分に主成分を用いるが、そのかわりに重心成分を使いたいときには、CENTROIDオプションを設定する。

```
proc varclus data=VCTEST CENTROID;
run;
```

Oblique Centroid Component Cluster Analysis				
4 Observations	PROPORTION =	0.75		
5 Variables	MAXEIGEN =	0		
Cluster summary for 1 cluster(s)				
Cluster	Members	Cluster Variation	Variation Explained	Proportion Explained
1	5	5.00000	2.39178	0.4784
Total variation explained = 2.391779 Proportion = 0.4784				
Cluster 1 will be split.				

解析結果の第1ページを示した。重心成分を使うときは第2固有値は計算されないためMAXEIGEN=の指定は無効となり、第1クラスター成分で75%の分散説明率というのが分割の基準となる。それ以外は、主成分の場合と同様に計算される。

### 5.2 クラスターの個数を制御する方法

最終的なクラスターの個数は、統計的基準よりも変数について研究者がもっている知識で決定するべきであろう。クラスターの個数を直接制御するには、MAXCLUSTERS=オプションを使用する。ここで指定された数値までクラスター分割が行なわれる。

```
proc varclus data=VCTEST MAXCLUSTERS=4;
run;
```

また、PROPORTION=またはMAXEIGEN=オプションを操作して、間接的に制御することもできる。

### 5.3 計算資源を節約する方法

大規模なデータを分析する場合は、計算時間と使用メモリの両面で制限ができる。あらかじめクラスターの個数がわかっている場合は、MINCLUSTERS=とMAXCLUSTERS=のオプションでクラスター数の下限と上限を設定しておく、計算時間が短くて済む。

```
proc varclus data=VCTEST MINCLUSTERS=3 MAXCLUSTERS=3;
run;
```

上の指定では、変数全体を一度に3個のクラスターに分割する。HIERARCHYオプションを指定して、階層型の制約をつけると計算時間

が短くなる。なぜなら、探索する候補が少なくなるからである。メモリを節約する方法は特になし。必要なメモリ量は、マニュアルでは、変数の数を $v$ 、クラスターの数を $c$ として、

$$v^2 + 2vc + 20v + 15c.$$

であると記載されているが、これは誤りであり、正しくは次のようになる。

$$190v + 144c + 8v^2 + 16vc$$

### 5.4 入力データの形式

今回の例では、多変量型のデータを入力形式とした。しかし、VARCLUSは相関行列または分散共分散行列から分析をスタートさせるものであるから、相関行列または分散共分散行列を直接入力データとすることができる。たとえば、次のようにする。

```
data VCTEST2(type=corr);
input _name_ $ x1-x5;
_type_ = 'corr';
cards;
X1 1.00000 0.89765 0.45796 0.49955 0.19260
X2 0.89765 1.00000 0.02268 0.10246 0.06201
X3 0.45796 0.02268 1.00000 0.87782 0.42056
X4 0.49955 0.10246 0.87782 1.00000 -0.05384
X5 0.19260 0.06201 0.42056 -0.05384 1.00000
;
proc varclus data=VCTEST;
run;
```

また、因子分析の結果を入力データとすることもできる。特に斜交回転の結果との比較に役に立つ。因子分析の結果は、PROC FACTORのOUTSTAT=オプションによりTYPE=FACTORのデータセット形式で出力されるので、これをPROC VARCLUSのDATA=オプションで受ける。

```
proc factor data=VCTEST rotate=PROMAX outstat=LOAD nfact=2;
run;

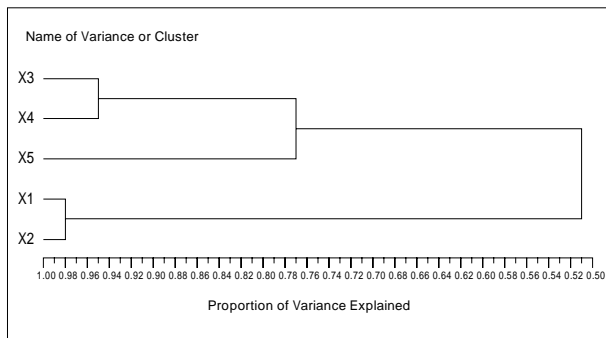
proc varclus data=LOAD;
run;
```

### 5.5 デンドログラム出力

VARCLUSの結果は、デンドログラムで表現することもできる。デンドログラムを描くためには、まずデンドログラムを描くのに必要なデータをPROC CLUSTERと同様にOUTTREE=オプションで指定したデータセットに出力させ、それをPROC TREEで受ける。最後まで分類を続けるため、変数の数と同じ数をMAXC=オプションで指定する。OUTTREE=オプションを指定すると、必然的にHIERARCHYオプションを指定したときと同様、階層型の制約の下で分割が行なわれる。VARCLUSによるデンドログラムの高さは説明された分散の割合を使うのが適切なので、OUTTREE=データセット中の\_PROPOR\_という変数をHEIGHTステートメントに指定する。変数の数が多いときには、\_NCL\_を使った方が見やすくなる。なお、リリース6.11からGRAPHICSオプションが評価版として使えるようになった。これによりグラフィカルなデンドログラムを得ることができる。

```
proc varclus data=VCTEST maxc=5 outtree=VCTREE;
run;
```

```
proc tree data=VCTREE horizontal graphics;
  height _PROPOR_;
run;
```

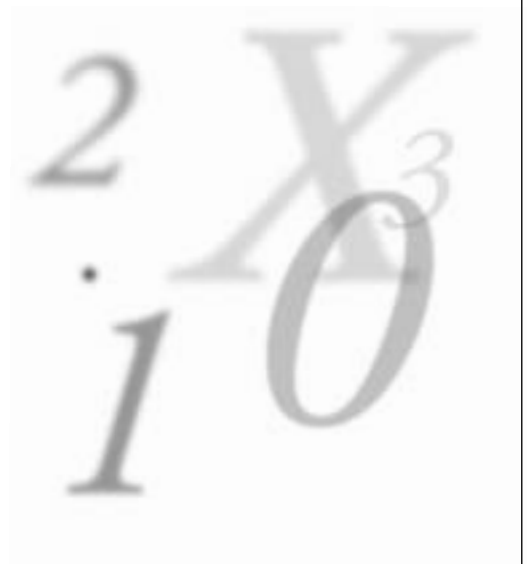


## 6. おわりに

多変量のデータを整理するとき、まず探索的因子分析を行なわれることが多い。しかし、探索的因子分析は技術的に難しい側面が存在する一方、現実のデータのモデルとしては不適切なことが多く、データ解析の手法としては問題が多いと著者は考えている。変数の整理・要約には、変数を直接クラスタリングするVARCLUSの方法が明快である。クラスター間の斜交性も自然に表現される。その後、データに対する知識が進めば、確認的因子分析、さらには共分散構造分析でデータの構造を表現するのが適切であろう。本稿をきっかけとして、VARCLUSという優れた手法が広く使われることを期待する。

### 参考文献

- 1) Asukai, N., Kato, H., Kawamura, N., Kim, Y., Yamamoto, K., Kishimoto, J., Miyake, Y., Nishizono-Maher, A. (2002) Reliability and validity of the Japanese-language version of the Impact of Event Scale-Revised (IES-R-J): Four studies on different traumatic events. The Journal of Nervous and Mental Disease.
- 2) Harris, C.W. and Kaiser, H.F. (1964), Oblique factor analytic solutions by orthogonal transformations, Psychometrika, 29, 347-362.
- 3) SAS Institute Inc. (1989), SAS/STAT User's Guide, Version 6, Forth Edition, Volume 2, Cary, NC: SAS Institute Inc.
- 4) 岸本 淳司 (1996) 変数のクラスタリング—PROC VARCLUS再発見—日本SASユーザー会論文集



# Q&A



全ての値が欠損値の数値変数と文字変数を削除する  
 各ユーザーごとに固有または任意のディレクトリをSASUSERライブラリとして割り当てる方法  
 SASシステムオプションの値をマクロ変数に格納する方法  
 データセットの存在判定  
 バッチ処理の際にパラメータを渡したい  
 標準入力からプログラムを読み込む方法  
 Windows版SASシステムリリース8.2で、データセットをMS Excelに出力する際に変数ラベルを出力する  
 SASシステムバージョン8におけるプログラムの自動保存  
 ODS HTMLステートメントでTABULATE プロシジャの出力幅を変更する

## Q

数値変数や文字変数の全ての値が欠損値の場合、それらの変数を削除する方法はありますか。

## A

プロシジャやDATAステップのオプションなどでは、そのような機能はサポートされていません。次のマクロで、全ての値が欠損値の変数を削除できます。

```

/* サンプルデータセットの作成 */
DATA sample;
  INPUT num1-num3 (char1-char4) ($);
CARDS;
1 . . A . . .
. . . . . B
3 . . C . . .
;
RUN;

%MACRO rm_miss(ds,out_ds);

*** ds: 入力データセット ***;
*** out_ds: 出力データセット ***;

/* 文字変数名を取得 */
PROC CONTENTS DATA=&ds OUT=_out1(WHERE=(type=2))
NOPRINT;
RUN;

DATA _null_;
  SET _out1;
  CALL SYMPUT('n_c',LEFT(_n_));
  CALL SYMPUT('c_var' || LEFT (_n_),name);
RUN;

/* 数値変数名を取得 */
PROC CONTENTS DATA=&ds OUT=_out2(WHERE=(type=1))
NOPRINT;
RUN;

```

```

DATA _null_;
  SET _out2;
  CALL SYMPUT('n_n',LEFT(_n_));
  CALL SYMPUT('n_var' || LEFT (_n_),name);

RUN;

/* 各文字変数の欠損値以外の件数をカウント */
PROC SQL NOPRINT;
  %DO i = 1 %TO &n_c;
    SELECT COUNT(&&c_var&i) INTO: c&i FROM &ds
    WHERE &&c_var&i NOT = ' ';
  %END;
QUIT;

/* 各数値変数の欠損値以外の件数をカウント */
PROC SQL NOPRINT;
  %DO i = 1 %TO &n_n;
    SELECT COUNT(&&n_var&i) INTO: n&i FROM &ds
    WHERE &&n_var&i NOT = . ;
  %END;
QUIT;

/* 全ての値が欠損値の変数を削除 */
DATA &out_ds;
  SET &ds;
  %DO i=1 %TO &n_c;
    %IF &&c&i = 0 %THEN %DO;
      DROP &&c_var&i;
    %END;
  %END;
  %DO i=1 %TO &n_n;
    %IF &&n&i = 0 %THEN %DO;
      DROP &&n_var&i;
    %END;
  %END;
RUN;

PROC DATASETS LIB=work NOLIST;
  DELETE _out1 _out2;
RUN;
QUIT;
%MEND rm_miss;

%rm_miss(sample,sample2); /* マクロの実行 */

```



**Q** SASUSERライブラリの変更は、SASシステムの環境設定ファイルSasv8.cfgファイルを編集することで可能ですが、SASUSERライブラリをユーザーごとに、ユーザー固有または任意のディレクトリに指定することは可能でしょうか。

たとえば、SASUSERライブラリを、  
 ・ユーザー user1 に関しては「D:¥SASFILES¥USER1¥sasuser」  
 ・ユーザー user2 に関しては「D:¥SASFILES¥USER2¥sasuser」として指定したいと考えています。

**A** Windowsの環境変数として、ユーザー固有のディレクトリ(「D:¥SASFILES¥USER1」など)を定義し、Sasv8.cfgファイル中で環境変数を参照させることによって、このような指定が可能となります。

設定作業は、一般ユーザーとして行なう作業と、Administrator権限を持つユーザーとして再度ログインしてから行なう作業に分かれます。

一般ユーザーとして行なう作業(ユーザー固有の環境変数を定義する) SASを使用する全てのユーザーについて行ないます。

1)ユーザー環境変数を定義します。コントロールパネルから「環境」タブを選択し、「(ユーザー名)のユーザー環境変数(U)」の下にある項目のうち、いずれかをアクティブにします。  
 例)TEMPなどを選択し、反転させる

2)変数(V)に任意の環境変数の名前(ここでは「SASENV」とします)を、「値(L)」にディレクトリ名を入力します。

例1)ユーザーID「USER1」として ログオンした場合	変数(V): SASENV 値(L): D:¥SASFILES¥USER1
例2)ユーザーID「USER2」として ログオンした場合	変数(V): SASENV 値(L): D:¥SASFILES¥USER2

3)設定(E)ボタンおよび「適用(A)」ボタンを押したあと、「OK」をクリックします。

Administrator権限を持つユーザーとして行なう作業(定義した環境変数を参照する)

1)環境設定ファイルSasv8.cfgを編集します。Sasv8.cfg ファイル中、「-SASUSER」の後で指定されているパスを編集します。

編集する行

```
-SASUSER "%SASENV¥sasuser"
```

編集例

変更前	/* Setup the default SAS System user profile folder */ -SASUSER "C:\winnt\profiles\user\personal\MySASFiles\sasuser."
変更後	/* Setup the default SAS System user profile folder */ -SASUSER "%SASENV¥sasuser" /* この部分を変更 */

上記の作業を終えたあと、一般ユーザーのIDで再ログインしてSASシステムを起動させると、SASUSERライブラリが次のように変更されます。

```
D:¥SASFILES¥(ユーザーID)¥sasuser
```

なお、この設定が正しく反映されているかどうかは、次の手順で確認できます。

- 1)SASを起動後、SASのエクスプローラ画面を表示させます。
- 2)SASUSERライブラリのアイコンを右クリックし、「プロパティ」を選択します。
- 3)SASUSERプロパティダイアログが表示されます。初期表示で「一般プロパティ」と表示されている部分(または下三角形で表示されている部分)をクリックします。
- 4)新たにSASUSERライブラリとして割り当てたディレクトリのパスが表示されていることを確認してください。

**Q** プログラム中で最後に作成したデータセット名は、SASシステムオプションの「\_LAST\_」に格納されますが、マクロ変数として「\_LAST\_」の値を使いたい場合にはどうしたらよいでしょうか。

**A** SASシステムオプションは、システムで用意されている「dictionaryテーブル」の「dictionary.options」に格納されています。これを利用することでマクロ変数に格納することができます。

サンプルプログラム

```
%MACRO lastdata;
%GLOBAL lastdata;
PROC SQL NOPRINT ;
SELECT setting INTO :lastdata FROM
dictionary.options
WHERE optname = '_LAST_' ;
QUIT;
%LET lastdata = %trim(&lastdata);
%MEND lastdata;

%lastdata ;

%PUT 最終作成データセット:&lastdata ;
```

プログラム中のそれぞれの変数の説明は次の通りです。

setting	実際のSASシステムオプションの値
optname	SASシステムオプションの各項目名

**Q** WORKライブラリの中に「Data1」というSASデータセットがあるかどうかを調べて、分岐処理を行ないたいのですが、どのようにしたらよいのでしょうか。

**A** データセットが存在するかどうかを判定するには、SCL関数であるEXIST関数を使用します。EXIST関数は、指定のデータセットが存在すればリターンコード「1」、存在しなければリターンコード「0」を返します。ただし、この関数はSCLプログラムの外では使用できないため、マクロ関数%SYSFUNCを使って呼び出します。

次のサンプルプログラムは、データセットが存在する場合にPRINTプロシジャを実行し、存在しなければ「データセットXXは存在しません」というメッセージをログに出力します。

```

/** テストデータの作成 **/
DATA data1;
  a=1;
RUN;

/** データセットの存在を判定するマクロ **/
%MACRO exist(lib,dsn);

  %IF %SYSFUNC(EXIST(&lib..&dsn)) %THEN
    %DO;
      PROC PRINT DATA=&lib..&dsn;

      RUN;
    %END;
  %ELSE %PUT ** データセット &dsn は存在しません **;
%MEND;

/** パラメータを指定してマクロを実行 **/
%exist(work,data1); /* 1番目はライブラリ名 */
                  /* 2番目はデータセット名 */

```

#### 実行結果

work.data1が存在する場合:

PRINTプロシジャが実行されます。

work.data1が存在しない場合:

SASログに「\*\* データセット data1 は存在しません\*\*」というメッセージが出力されます。

## Q

Windows版SASシステムでバッチ処理を行なう際に、ユーザー固有の値をセットし、SASプログラム内で値を受け取って実行することは可能でしょうか。

## A

バッチプログラム内にユーザー固有のパラメータを記述してSASプログラムへ受け渡すには、「SYSPARM」システムオプションを使用します。SYSPARMオプションの引数には、最大200バイトまでの文字列を指定できます。SYSPARM関数を使用すれば、SASのDATAステップ内でこの引数を参照できます。その他、任意の場所で参照するには、自動マクロ変数「&SYSPARM」を使用します。以下に、バッチプログラムのサンプルを示します。拡張子「.bat」をつけて、任意のファイル名で保存します(例:test.bat)。なお、SYSPARMには、「JPNSAS」をセットしています。

```

"C:\Program Files\SAS
Institute\SAS\V8\nls\ja\sas.exe"
-sysin "C:\temp\test.sas"
-config "C:\Program Files\SAS
Institute\SAS\V8\nls\ja\Sasv8.cfg"
-log "C:\temp\test.log"
-nosplash
-icon
-sysparm JPNSAS

```

実際は1行で記述します。

バッチファイル内で参照するSASプログラム"test.sas"の内容です。

-sysparmで与えたパラメータをログに出力するプログラムとなっています。

```

/** test.sas **/
%PUT "UserIDは &sysparm です";

```

バッチファイルを実行すると、-logで指定した場所にSASシステムのログファイル「test.log」が作成されます。ログファイルの内容は、次のようになります(一部抜粋)。

```

NOTE: SAS システム初期化 処理 :
      処理時間           0.67 秒
      CPU 時間           0.42 秒

1           %PUT UserIDは &sysparm です;

UserIDは JPNSAS です

```

## Q

UNIX版のSASシステムで、標準入力から直接プログラムを読み込んで実行する方法はありますか。

## A

STDIOシステムオプションを指定することによって可能です。このオプションによって、プログラムを標準入力から読み込み、プロシジャ出力を標準出力に、ログを標準エラーにそれぞれ書き出します。次の例では、シェルプロンプトから以下の1行を実行することで、SETINITプロシジャの結果を出力させています。

```
$ echo proc setinit.; | sas -stdio
```

また、この機能によって、シェルスクリプトの中にSASプログラムを記述したり、SASの出力を他のコマンドの標準入力として渡すことが可能となります。次の例では、OPTIONSプロシジャと%PUTステートメントの出力から任意の文字列を検索するシェルスクリプトを作成し、実行しています。

```

$ cat stdio.sh
#!/bin/sh
(
  /usr/bin/sas -stdio 2>&1 << SRC
proc options;
run;
%put _automatic_;
SRC
) | .
grep $1
$ ./stdio.sh TIME
NOFULLSTIMER システムパフォーマンス統計量のログ出力を無効にします
STIMEFMT=M STIMER 出力で用いる出力形式を指定します
STIMER システムパフォーマンス統計量のログ出力を有効にします
AUTOMATIC SYSTIME 12:34

```

**Q** Windows版SASシステム リリース8.2で、SASデータセットをMicrosoft Excel(以下 Excel)に出力しているのですが、Excelの先頭行に変数ラベルを出力することができません。SASシステム リリース6.12ではこの機能を設定するオプションが、書出しウィザードにあったのですが、どのようにすればできますか。

**A** SASシステム バージョン8から、「ファイル」>「データのエキスポート」においてラベルの書出しができなくなりました。変数ラベルをExcelに出力するには、下記の2つの方法があります。

## 1) リリース6.12の「書出し」で使用していたDBLOADプロシジャを実行する

この方法の利点と欠点

利点	・SASデータセットのラベルを従来通り出力できる
欠点	・Excel 97および2000形式のファイルは出力できない ・ウィザードではなく、プログラムを作成して実行する ・バージョン8の拡張点(変数名の拡張、1変数あたりの最大文字数の拡張、および変数ラベルの長さ拡張等)を使用できない

### DBLOADプロシジャのサンプルプログラム

```
/* データセット名 */
PROC DBLOAD DBMS=EXCEL DATA=SASUSER.CLASS ;
  PATH='c:\temp\class.xls'; /* ファイル名 */
  PUTNAMES YES;
  LIMIT=0;
  LABEL;
  RESET ALL;
  LOAD;
RUN;
```

## 2) DDE(Dynamic Data Exchange)を利用する

この方法の利点と欠点

利点	・SASデータセットのラベルを従来通り出力できる ・バージョン8の拡張点を使用できる ・Excel 97/2000形式のファイルも出力できる
欠点	・ウィザードではなく、プログラムを作成して実行する ・Excelを起動するため、PCのリソースを必要とする

DDEを利用する場合は、実行するプログラムを作成する必要があります。参考のため、Base SASソフトウェアのSASマクロ機能を利用するサンプルプログラムを以下に示します。

注:下記の例ではExcel 2000を使用しています。

```
%MACRO _sas2xls(dsn,xlsname);

OPTIONS NOXWAIT NOXSYNC;
x "start excel"; /* Excelを起動 */

DATA _null_;
  x=SLEEP(5); /* Excelの起動を待つためにSASをSLEEP */
RUN;
```

```
PROC CONTENTS DATA=&dsn
  OUT=_ddel(KEEP=name varnum label nob);
NOPRINT;
RUN;

PROC SORT DATA=_ddel OUT=_ddel(DROP=varnum);
  BY varnum;
RUN;

DATA _ddel;
  SET _ddel END=last;
  IF label=" " THEN label=name;
  /* マクロ変数 _name1~_nameN に変数名を格納 */
  CALL SYMPUT('_name' || LEFT(PUT(_n_,8.)),
    LEFT(COMPRESS(name)));
  /* マクロ変数 _lbl1~_lblN にラベルを格納 */
  CALL SYMPUT('_lbl' || LEFT(PUT(_n_,8.)),
    LEFT(COMPRESS(label)));
  IF last=1 THEN DO;
    /* マクロ変数 _last にOBS数を格納 */
    CALL SYMPUT('_last',LEFT(COMPRESS(PUT(_N_,8.))));
    /* マクロ変数 _obs に元データセットのOBS数を格納 */
    CALL SYMPUT('_obs',LEFT(PUT(nobs,8.)));
  END;
RUN;

%LET _ddeobs=%EVAL(&_obs+1);
FILENAME export dde
"EXCEL|Sheet1!r1c1:r&_ddeobs.c&_last" NOTAB;

DATA _null_;
  SET &dsn;
  FILE export;
  IF _n_ = 1 THEN
    PUT %DO num=1 %TO &_last; /* ラベルを出力 */
    "&_lbl&num"

    %IF &num ne &_last %THEN %DO;
      '09'x
    %END;
  %END;
  ;

  PUT %DO num=1 %TO &_last; /* 変数値を出力 */
    &&_name&num
    %IF &num ne &_last %THEN %DO;
      '09'x
    %END;
  %END;
  ;
RUN;

OPTIONS NOXWAIT XSYNC;
FILENAME data dde 'excel|system';

DATA _null_; /* シートに名前を付けて保存し、Excelを終了 */
```

```

FILE data;
PUT "[SAVE.AS("&xlsname")]";
PUT "[QUIT()]";
RUN;

OPTIONS XWAIT XSYNC;

%MEND _sas2xls;

/* 実際の利用 %_sas2xls(データセット,Excelのファイル名) ; */
%_sas2xls(sashelp.prdsale,c:\temp\test.xls);

```

**Q** SASシステムリリース8.2を利用中、プログラムの「自動保存を完了しました」というメッセージが表示されます。実際にプログラムはどのフォルダに自動保存されていますか。また、設定はどのようにすればよいですか。

**A** SASシステムバージョン8から、拡張エディタ、プログラムエディタの内容が自動的に保存される機能が追加されました。実際に保存されるフォルダとファイル名、および設定方法(機能の有効/無効、自動保存の間隔など)は以下のようになります。

#### Windows版

拡張エディタの保存	Windowsで指定される一時フォルダ内(環境変数名 %temp% に指定されるフォルダ)に、「Autosave of 無題.SAS」などのファイル名で保存される
プログラムエディタの保存	現在の作業フォルダに「pgm.asv」として保存される
設定方法	ツールバーの「ツール」「オプション」「プリファレンス」「プリファレンス」ダイアログボックスの[編集]タブ

#### UNIX版

拡張エディタ、プログラムエディタの保存	現在の作業フォルダに「pgm.asv」として保存される
設定方法	ツールバーの「ツール」「オプション」「プリファレンス」「プリファレンス」ダイアログボックスの[DMS]タブ

**Q** TABULATEプロシジャでの出力を、ODS HTMLステートメントを使用してHTML形式で出力しています。その際に、TABULATEプロシジャで出力される表全体の幅を広げるにはどうすればよいでしょうか。

**A** ODS HTMLステートメントの出力で、TABULATEプロシジャを使用して出力される表全体の幅は、TABLEステートメントのSTYLE属性を指定することで調整できます。右記のプログラム例をご覧ください。

```

ODS HTML FILE='c:\temp\test.html' ;
PROC TABULATE DATA=sashelp.class;
  CLASS sex age;
  VAR weight height;
                                     /* 指定なし */
  TABLE sex*age, (weight height)*mean;
                                     /* 10cm の指定 */
  TABLE sex*age, (weight height)*mean /
  STYLE=[OUTPUTWIDTH= 10cm] ;
RUN;
ODS HTML CLOSE;

```

OUTPUTWIDTHスタイル属性は、下記の単位で指定可能です。

CM	センチメートル
MM	ミリメートル
CM	インチ
MM	ピクセル数
%	パーセント(ブラウザ幅に対する比率であり、本来の幅に対する比率ではありません。)

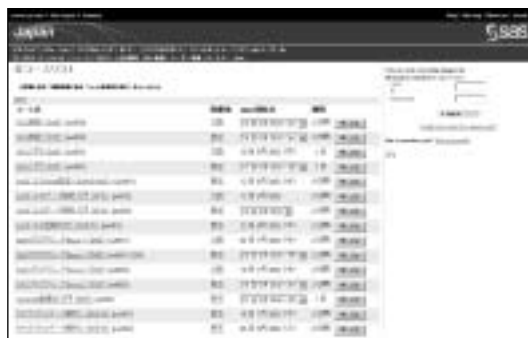
TABULATEステートメントで使用可能なスタイルについては、オンラインヘルプのTABULATEプロシジャの項目をご参照ください。また、TABULATEプロシジャのスタイル属性で指定可能な属性の詳細については、「SAS Procedures Guide, Version 8, Volume 1 and 2」(注文番号:57238)の「Procedure Concepts」の項目でご確認いただけます。

# SAS Training

## SASトレーニングのお知らせ

### オンライン申し込み開始

このたび、弊社ではオンラインによるトレーニング受付を開始しました。従来のFAXによるお申し込み以外に、インターネット経由で簡単に申し込めるようになりましたので、どうぞご利用ください。



URL: <http://www.sas.com/japan/training/>

### セットプランのご案内

前号のSAS Technical News(Spring 2002)でご案内しましたセットプラン「Statistic GUI Plan」は、都合により「Statistic Plan1」および「Statistic Plan2」へ変更させていただいております。急な変更によりお客さまにはご迷惑をお掛けしましたことお詫び申し上げます。

Statistic Plan 1	<3コースセット：¥95,000(税別)>
<ul style="list-style-type: none"> <li>・「統計概論」コース</li> <li>・「SASによるデータ解析入門」コース</li> <li>・以下3コースよりいずれか1コース選択                     <ul style="list-style-type: none"> <li>「主成分および因子分析」コース</li> <li>「分散分析」コース</li> <li>「判別およびクラスター分析」コース</li> </ul> </li> </ul>	

Statistic Plan 2	<3コースセット：¥110,000(税別)>
<ul style="list-style-type: none"> <li>・「統計概論」コース</li> <li>・「SASによるデータ解析入門」コース</li> <li>・以下3コースよりいずれか1コース選択                     <ul style="list-style-type: none"> <li>「SASによる回帰分析」コース</li> <li>「カテゴリカルデータ解析1」コース</li> <li>「カテゴリカルデータ解析2」コース</li> </ul> </li> </ul>	

なお、「Basic Plan」も継続して受け付けております。

Basic Plan	<3コースセット：¥100,000(税別)>
<ul style="list-style-type: none"> <li>・「SASプログラミング: Basic I」コース</li> <li>・「SASプログラミング: Basic II」コース</li> <li>・以下5コースよりいずれか1コース選択                     <ul style="list-style-type: none"> <li>「マクロ言語入門」コース</li> <li>「マクロ言語上級」コース</li> <li>「ODS入門」コース</li> <li>「実践データハンドリング」コース</li> <li>「Version 8拡張点入門」コース</li> </ul> </li> </ul>	

有効期限は1コース目初日より6ヶ月間となります。  
 請求書は、1コース目受講終了後に発行いたします。  
 日程・コースの変更は可能ですが、キャンセルはできませんのでご了承ください。

### [セットプランのお申込方法]

弊社指定のトレーニング申込書のセットプランの欄をチェックした上で、受講希望コースと日程を明記し、FAXでお送りください。

### 「カテゴリカルデータ解析2」コース開催日決定

トレーニング内容改訂のために、開催を中止しておりました「カテゴリカルデータ解析2」コースの日程が下記のとおり決定いたしました。

開催日: 8月6・7日(火・水)  
 開催場所: 東京会場

その他のコース情報については、弊社ホームページ <http://www.sas.com/japan/training/> をご参照いただくか、下記<トレーニング担当>までお問い合わせください。

トレーニング担当
T E L 03-3533-3835
F A X 03-3533-3781
E-mail <a href="mailto:JPNTraining@sas.com">JPNTraining@sas.com</a>
お問い合わせメールアドレスが変わりましたのでご注意ください。



# Programmers Workshop

## SASプログラマーズ・ ワークショップ開催のお知らせ

今回のSASプログラマーズ・ワークショップでは、Enterprise Guideソフトウェアの最新リリースである2.0をご紹介します。今までお使いいただいているリリース1.2からの変更点やセットアップ方法などを中心にお伝えするとともに、この製品についてお寄せいただいた過去のご質問より、最新リリースでもお使いいただける情報について併せてご紹介します。

日程：9月12日(木)

会場：株式会社SASインスティテュートジャパン 東京本社  
プレゼンテーションルーム

住所：東京都中央区勝どき1-13-1 イヌイビル・カチドキ8F

時間：15:00～17:00

定員：50名

日程：9月27日(金)

会場：株式会社SASインスティテュートジャパン 大阪支店  
プレゼンテーションルーム

住所：大阪市北区堂島浜1-4-16 アクア堂島西館12F

時間：15:00～17:00

定員：50名

両日とも参加は無料となりますので、ぜひご参加ください。なお、定員に限りがありますので、お早めにお申し込みください。お申し込み方法は、「SASプログラマーズ・ワークショップ参加申込用紙」へご記入の上、FAX、E-mailにてご送付ください。SASプログラマーズワークショップのお申し込み用紙は下記のURLより入手できます。

<http://www.sas.com/japan/service/technical/pws.html>

その他、ご不明点、ご質問がございましたら、以下までご連絡ください。

### プログラマーズ・ワークショップ担当

T E L 03-3533-3835

F A X 03-3533-3781

E-mail [pws@jpn.sas.com](mailto:pws@jpn.sas.com)

# New Publications

## 新刊マニュアルのご紹介

「Windows版 SASシステム 使用の手引き バージョン8」

注文番号：17171

価 格：13,000円(税抜)

本書は、Windows版 SASシステム バージョン8が提供する、Windowsオペレーティングシステムに依存する機能をまとめたマニュアルです。本書には、Windows環境におけるSASシステムの起動から各種機能の使用法、Windows版 SASシステムに特有のグラフィカルユーザーインターフェイス、ステートメント、関数、オプションなどが詳細に解説されています。Windows版 SASシステムを初めてお使いになる方はもちろん、Windows環境に特化したSASシステムの機能を使用する方はぜひ本書をお読みください。Windows版 SASシステム リリース8.2を使用する場合、「What's New in SAS Software for Release 8.1」と「What's New in SAS Software for Release 8.2」も併せて参考にしてください。

なお、上記のマニュアルは、SASシステム リリース 8.2 のインストールキットに含まれている「インストレーションドキュメント メディア」(CD-ROM)に収録されています。このインストールキットは、SASシステムリリース8.2をご契約されている全サイトに無償でお送りしています(1契約につき1枚)。

### SASロゴ入りグッズ販売のご案内

2002年4月より、SASロゴ入りグッズの販売を開始しました。弊社ホームページより商品の詳細をご覧いただけます。SASグッズは、マウスパッド等多数取り揃えておりますので、ご興味のある方はぜひご利用ください。お申し込み方法はマニュアルの注文と同様です。

<http://www.sas.com/japan/manual/goods.html>

### マニュアル販売係

T E L 03-3533-3835

F A X 03-3533-3781

E-mail [JPNBooksale@sas.com](mailto:JPNBooksale@sas.com)

お問い合わせメールアドレスが変わりましたのでご注意ください。

なお、マニュアル申込用紙および最新のPublication Catalog(マニュアル案内パンフレット)は、弊社ホームページ(<http://www.sas.com/japan/manual/>)にて公開していますので、併せてご利用ください。



# Web Based Training

## インターネット新時代に SAS e-Learning 始動

SASでは新教育マテリアルとして、Web Based Training(以下WBT)を2002年9月1日より開始します。

今までは・・・

従来、SASを習得するためには、日時・会場の決まったトレーニングに参加する必要がありました。そのため、トレーニング受講を希望しても、スケジュールや会場の制約により参加できない場合がありました。

これからは・・・

今回ご紹介するWBTは、従来のトレーニングとは異なり、インターネットを利用できる環境さえあればいつでも自由に受講することが可能です。このため、期間内であれば時間に制約されずに学習することができます。また繰り返し学習することで、より理解を深めることができます。社内での研修に使用する場合にも、トレーニングルームの確保や受講者のスケジュール調整が必要ありません。

WBTの特長

機能・操作方法、およびSASを使う上での注意点は、章ごとに音声で説明します。また各画面上でマウスカーソルが自動的に動くため、必要な操作を目で確かめることができます。各画面には「巻戻し」「繰返し」「一時中止」のボタンがあるので、各自の学習速度を自由に調節できます。さらに、各章末には、WBTを効果的に利用するための三択式の簡単なクイズを用意しています。

受講価格

- ・1人受講につき 20,000円(消費税別)
  - ・アカデミック料金 8,000円(消費税別)
- (大学生、大学院生対象：申込書に学生証のコピーを添えてお送りください)

まとめて50名以上でのお申し込みには、別途バリュー価格を適用します。詳細についてはお問合せください。

コース内容の詳細、および受講申し込み方法

1. 下記URLにアクセスして、WebまたはFAXにてお申し込みください。  
<http://www.sas.com/japan/training/>  
個人単位でお申し込みください。
2. お申し込みいただいた後、請求書を送付します。入金確認後、ご利用に必要なIDとパスワードをE-mailにて送付します。なお、受講者数は1ヶ月に50名様までとさせていただきますので、お早めにお申し込みください。

WBT受講期間について

専用のIDとパスワードが送付されてから、最大で2ヶ月間受講できます。2ヶ月以内で先行して終了することも可能です。受講期間が2ヶ月を超えた場合は、自動的にトレーニングは中断となります。再度受講を希望する場合は、新たにお申し込みが必要となりますのでご注意ください。トレーニング終了後、株式会社SASインスティテュートジャパンより「コ

ース終了証」を送付します。なお、WBTの進捗によっては、WBT管理者から確認のメールが送付されます。

WBT動作環境

受講には以下のインターネット環境でのご利用を推奨します。

回線速度 : ISDN 64Kbps以上

ソフトウェア : OS Microsoft Windows 98(SE)以上

ブラウザ Microsoft Internet Explorer 5.01 SP2以上

お問合せ先

株式会社SASインスティテュートジャパン エデュケーショングループ

T E L 03-3533-3835

F A X 03-3533-3781

E-mail JPNTraining@sas.com

<http://www.sas.com/japan/training/>

なお、コースにより受講価格や期間が変更される場合がありますので、お申し込みの際は、弊社ホームページでご確認ください。



# Latest Releases

## 最新リリース情報

### PCプラットフォーム

Windows版	SASシステム リリース8.2	TS2M0
OS/2版	SASシステム リリース6.12	TS020
Macintosh版	SASシステム リリース6.12	TS040

### ミニコンピュータプラットフォーム

OpenVMS AXP版	SASシステム リリース6.12	TS020
OpenVMS VAX版	SASシステム リリース6.08	TS407

### UNIXプラットフォーム

MIPS ABI版	SASシステム リリース6.11	TS040
Tru64版	SASシステム リリース8.2	TS2M0
ABI+版	SASシステム リリース6.11	TS040
SunOS/Solaris版	SASシステム リリース8.2	TS2M0
HP-UX版	SASシステム リリース8.2	TS2M0
AIX版	SASシステム リリース8.2	TS2M0
Linux版	SASシステム リリース8.2	TS2M0

### メインフレームプラットフォーム

MVS版	SASシステム リリース6.09E	TS470
MSP版	SASシステム リリース6.09E	TS470
VOS3版	SASシステム リリース6.09E	TS470
CMS版	SASシステム リリース6.08	TS410

## SAS Technical News Summer 2002

発行  
株式会社SASインスティテュートジャパン

テクニカルニュースに関するお問い合わせ先

テクニカルサポートグループ

TEL: 03-3533-3877

FAX: 03-3533-3781

E-mail: technews@jpn.sas.com



株式会社SASインスティテュートジャパン

[www.sas.com/japan/](http://www.sas.com/japan/)

東京本社  
〒104-0054  
東京都中央区勝どき1-13-1  
イヌイビル・カチドキ 8F  
Tel 03 (3533) 6921  
Fax 03 (3533) 6927

大阪支店  
〒530-0004  
大阪市北区堂島浜1-4-16  
アクア堂島西館 12F  
Tel 06 (6345) 5700  
Fax 06 (6345) 5655