

# SAS Technical News

Winter 2009

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

## CONTENTS

- 1 SAS 9.2への移行について
- 5 SAS Academic News
- 7 Q&A
- 14 SASトレーニングのお知らせ
- 16 マニュアル販売に関するお知らせ
- 16 最新リリース情報
- 16 SAS Technical News送付についてのご案内

## 特集

### SAS 9.2への 移行について

Foundation  
SAS® 9

#### 1. はじめに

SASの最新リリースであるSAS 9.2がリリースされ、3カ月あまり経ちました。今号の特集では、移行を検討されるお客様を対象に、SAS 9.2 Foundationへの移行の概要、方法論、注意点などをご紹介します。具体的な移行方法については、次号の特集にて掲載予定です。なお、SAS 9.2の新機能に関しては、前回のSAS Technical News Autumn 2008の「特集1 SAS 9.2のご紹介(Foundation編)」にて掲載しておりますので、ご確認ください。

SAS Technical News Autumn 2008

<http://www.sas.com/japan/periodicals/technews/pdf/08aut.pdf>

#### 2. 移行の概要

SAS 9.2への移行は、計画の立案、環境の構築、計画に基づいた移行作業の3つのフェーズで行ないます。

- 1) 計画(移行前の準備:移行計画立案)
  - ・ドキュメントの入手方法(インストールセンターについて)
  - ・SAS Foundationシステム必要条件について
  - ・既存資産のボリュームの確認
  - ・既存資産の分類
  - ・移行ゴールの設定
- 2) 環境構築(新環境の構築)
  - ・SAS 9.2のインストール
  - ・SAS 9.2の動作検証ツール
  - ・複数バージョンのSASを共存させる場合の注意点
  - ・Windows版SAS 9.2固有の注意点
  - ・プロダクト別の注意点

3) 移行の環境準備および移行作業

- ・ SASシステムオプション
- ・ 自動実行ファイル(AUTOEXEC)
- ・ SASライブラリ(データセット、ビュー、カタログ)の移行
- ・ その他の注意点
- ・ 新環境への切り替え

3. 計画(移行前の準備:移行計画立案)

移行計画の手始めとして、移行先環境または、移行のためのテスト環境のオペレーティングシステム(以下、OSと略)が、SAS 9.2のシステム必要条件を満たしているかを確認します。次に、現在ご利用中のプロダクトに依存したシステム環境のバージョン、例えば、データベースやJREなどのバージョンの確認を行ないます。

SAS/ACCESSについて

外部データソースとしてデータベースをご利用の場合、DBMS製品のソフトウェアのバージョンがシステム必要条件を満たしているかどうか確認します。システム必要条件に記載されていないバージョンをご利用の場合は、弊社テクニカルサポートまでお問い合わせください。

JRE(Java Runtime Environment)について

SAS 9.2 FoundationはJRE 5を必要とします。システム必要条件に記載されていないバージョンをご利用の場合は、弊社テクニカルサポートまでお問い合わせください。

3.1 ドキュメントの入手方法(インストールセンターについて)

SAS 9.2のインストールに関連するドキュメントは、弊社テクニカルサポートのWebサイトのインストールセンターから入手してください。掲載されているドキュメントには、以下の種類があります。

- ・ SAS Foundationシステム必要条件
- ・ プロダクトシステム必要条件
- ・ クイックスタートガイド
- ・ SAS Deployment Wizardユーザーガイド
- ・ SAS Foundation使用上の注意点
- ・ SAS検証ツールユーザーガイド
- ・ SAS Foundation設定ガイド
- ・ SAS Foundationソフトウェアの更新手順

インストールセンターのURL

<http://www.sas.com/japan/service/documentation/installcenter/index.html>

なお、インストールセンターおよびSASインストールに関するドキュメントの概要は、前号の「SAS Technical News Autumn 2008」の「特集2 SAS 9.2対応インストールセンターのご紹介」にて掲載しておりますので、あわせて参考にしてください。

3.2 SAS Foundationシステム必要条件について

SAS Foundationシステム必要条件には、SAS 9.2がサポートするOS、CPUの種類、メモリ容量、使用ハードディスク容量などの稼働要件が掲載されています。このドキュメントを確認し、SAS 9.2の稼働要件を満たしている移行先環境を構築する必要があります。ただし、SAS 9.2では、一部サポート対象外となったOSもありますので、ご注意ください。

システム必要条件に記載されているメモリ・ハードディスク容量は、必要最小限の値です。このため、ご利用形態によっては、より多くのメモリ・ハードディスクを搭載する必要があります。現状搭載されているメモリ・ハードディスクの容量を基準として、現在の運用におけるパフォーマンスデータやジョブの実行状況をもとに、メモリ・ハードディスク容量の見直しをご検討ください。

3.3 既存資産のボリュームの確認

既存資産(SASプログラム、データセット、カタログなど)の棚卸のために、ボリューム確認を行ないます。具体的には、SASの拡張子(SAS 8からの移行の場合は、.sas7bdat、.sas7bcatなどのファイル)をシステム全体から検索し、その容量を特定します。また、SASプログラムをスケジュール化して実行している場合には、ジョブの本数、実行頻度もあわせて確認します。なお、各エンドユーザー部門が多岐にわたる場合は、既存資産のボリュームや実行頻度などを、インタビューや質問表などで確認するのが一般的です。

3.4 既存資産の分類

重要度、実行頻度をもとに、既存資産を1~4のカテゴリに分類します。以下図1の縦軸は「実行頻度」、横軸は「重要度」を表しています。

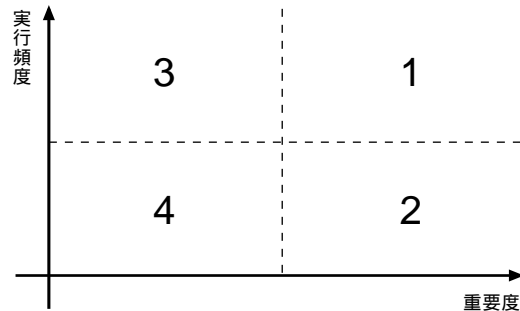


図1:プログラム資産の分類図

上記分類をもとに、移行テストに使用するSASアプリケーションを選定します。既存資産のボリュームが少ない場合、上記のカテゴリに分類することなく、全てのアプリケーションを実行し検証していただく方法があります。しかしながら、日常的に業務アプリケーションとしてSASを使用している場合、ある一定数のプログラムが存在するのが一般的です。これら全てのプログラムを実行し検証することが難しい場合、カテゴリ毎に実行本数をサンプリングすることにより、効率的に検証することができます。

1のカテゴリでは、重要度および実行頻度も高いため、基本的には全てのアプリケーションの検証が望ましいと考えます。しかしながら、全てのアプリケーションの実行による検証が難しい場合、代表的なアプリケーションからサンプリングし、検証を行なう方法があります。

該当するアプリケーションとしては、日時の夜間バッチなどで起動されるジョブや週次、月次そして四半期など、スケジュールに登録され定期的に行うバッチジョブ、業務上、日常的に実行しているアプリケーションなどが想定されます。

2のカテゴリは、上記1のカテゴリより実行頻度は少ないものの重要度が高いため、基本的には1のカテゴリ同様、全てのアプリケーションを実行し、検証することが望ましいと考えます。ただし、全ての実行と検証が難しい場合、やはりサンプリングによるアプリケーションの選定にて、検証を行なう方法があります。

3のカテゴリは、アプリケーションの実行頻度は高いものの重要度が低いため、代表的なアプリケーションのサンプリングのみでの検証で十分であると想定されます。サンプリングの本数自体も、カテゴリ1および2より少ない本数とし、かつ、アプリケーションの処理内容を特定し、検証する方法があります。

4のカテゴリは、実行頻度が少なく、重要度も低いため、最小限のサンプリングのみでの検証で十分であると想定されます。

### 3.5 移行ゴールの設定

移行作業終了の判断をするために、ゴールを設定します。その目安としては、以下のようなものが挙げられます。

- ・既存のアプリケーションがエラー出力せずに正常に動作する
- ・既存のアプリケーションの処理時間が旧SASバージョンと同等である
- ・既存のアプリケーションの計算結果レポートのレイアウトが旧SASバージョンと同等である

## 4 環境構築(新環境の構築)

SAS 9.2のシステム必要条件に沿って構築した環境へ、SAS 9.2をインストールします。また、本章では旧バージョンのSASとSAS 9.2を並行稼働させる場合の注意点もあわせてご紹介します。

### 4.1 SAS 9.2のインストール

SAS Deployment Wizardユーザーガイドを参照し、SAS 9.2のインストールを行います。なお、SAS 9.2よりインストールメディアが従来のCD-ROMからDVDに変更されましたので、ご注意ください。

### 4.2 SAS 9.2の動作検証ツール

SAS 9.2では、インストールが正常に行なわれたかを検証するツールを、標準で用意しています。

- ・SAS Installation Qualification Tool (以下、SAS IQと略)

この検証ツールを実行すると、その検証結果がPDF、HTML形式ファイルで出力されますので、導入に関する情報の一部として保存しておくことをお勧めいたします。

以下、SAS IQの実行方法について説明いたします。

#### SAS IQの実行手順

SAS IQ は、以下の手順で実行します。

#### ・Windowsの場合

Windowsの[スタート]メニューから、[すべてのプログラム] - [SAS] - [SAS 9.2 License Renewal & Utilities] - [SAS Installation Qualification Tool]を選択します。

#### ・UNIXの場合

SAS 9.2のインストールユーザーにてログインし、以下のコマンドを実行します。

```
<SAS 9.2のインストールディレクトリ>/sastest/sasiq.sh
```

- 1) 出力したいディレクトリ名を入力します。この例では、C:%tempを指定しています。

```
Please specify a directory where the output files should go [.] : C:%temp
```

- 2) 指定したディレクトリに、以下の名前のファイルが作成されていることを確認します。

```
C:%temp%validate_yyyyymmdd.xxx%sasiq.htm
C:%temp%validate_yyyyymmdd.xxx%sasiq.pdf
```

- 3) 手順2)で作成されたファイルを確認し、「Passed」メッセージが出力されていることを確認します。「Failed」メッセージが出力されている場合は、インストールに問題がある可能性が高いため、該当するプロダクトを再インストールしてください。



図2:SAS IQの出力例

なお、SAS 9.2検証ツールのユーザーガイドは、以下の場所にあります。

SAS 9.2検証ツール ユーザーガイド

[http://www.sas.com/japan/service/documentation/installcenter/instqualtool/9.2/qualification\\_tools\\_guide.pdf](http://www.sas.com/japan/service/documentation/installcenter/instqualtool/9.2/qualification_tools_guide.pdf)

ご不明点に関しては、弊社テクニカルサポートまでご連絡ください。

### 4.3 複数バージョンのSASを共存させる場合の注意点

移行を前提として、1台のマシンに旧バージョンのSASと、SAS 9.2を並行稼働させることが可能です。なお並行稼働させる場合、一部手続きが必要になりますので、弊社担当営業までご連絡ください。以下に並行稼働の際の注意点をご紹介します。

#### インストールディレクトリ

各々のSASは、異なるディレクトリへインストールしてください。

#### WORKライブラリ

各々のSASにて使用するWORKライブラリのディレクトリは、異なるディレクトリを使用してください。WORKライブラリとして使用するディレクトリのパスの設定は、環境設定ファイル中のWORKシステムオプションにて変更可能です。

#### ・環境設定ファイル名

UNIX版SAS 9.2 : <!SASROOT>/sasv9.cfg

Windows版SAS 9.2 : <!SASROOT>%nls%ja%sasv9.cfg

#### ・WORKシステムオプションの変更例(Windows版)

-WORK "D:%SASWORK92"

#### ・WORKシステムオプションの変更例(UNIX版)

-WORK "/tmp/SASWORK92"

なお、起動時オプションにて、WORKライブラリを変更されている場合もありますので、設定内容に関してご注意ください。

#### SASUSER PROFILEカタログ

SASUSER PROFILEカタログは、異なるバージョンにて流用することができません。SAS 9.2にて旧バージョンと同様に、SASUSER.PROFILEカタログをカスタマイズして使用する場合は、再度設定を行なってください。

#### 4.4 Windows版SAS 9.2固有の注意点

旧バージョンのWindows版SASと、Windows版SAS 9.2を並行稼働させる場合の注意点を紹介します。

##### 複数バージョンのSASのインストール順序

移行のためのテスト環境を構築する際、新たに複数バージョンのSASをインストールする場合があります。必ず古いバージョンのSASからインストールする必要があります。

例えば、SAS 8.2とSAS 9.2を共存させる場合、SAS 8.2、SAS 9.2の順序でインストール作業を行います。

SASはSAS Shared Filesと呼ばれるコンポーネントを「Shared Files」フォルダへ保存し、バージョン間で共有します。先に新しいバージョンのSASをインストールすると、後からインストールした旧バージョンのコンポーネントが上書きされてしまい、新しいバージョンのSASが動作しなくなります。

##### ファイル拡張子の関連付け

旧バージョンのSASに拡張子の関連付けを変更する場合、インストールメディア中のsetup.exeを実行して、ウィザードに従って作業を行います。詳細に関しては、弊社FAQサイトをご参照ください。

##### 拡張子の関連付けを再実行する方法

<http://www.sas.com/japan/service/technical/faq/list/body/inw018.html>

#### 4.5 プロダクト別の注意点

旧バージョンのSASとSAS 9.2を並行稼働させる場合の各プロダクトの注意点を紹介します。

##### SAS Integration Technologies

ロードバランサー（もしくは依存するスポーナー）、オペレータポート、コミュニケーションポートをバージョンごとに分ける必要があります。オブジェクトスポーナーの環境設定ファイルは、バージョンごとに別にしてください。使用するポート番号については、ネットワーク管理者の方へご確認ください。

##### SAS/CONNECT、SAS/SHARE

SAS/CONNECTのスポーナー、SAS/SHAREサーバーのポート番号は、バージョンごとに分ける必要があります。使用するポート番号については、ネットワーク管理者へご相談ください。Windows版にてスポーナーをサービス起動する場合は、起動ユーザーを登録する必要があるため、各セキュリティポリシーに沿ってユーザーを登録してください。

##### SAS/IntrNet

旧バージョンのSAS/IntrNetとSAS/IntrNet 9.2は1台のマシンに共存させて使用させることができますが、以下の注意点があります。

- ・アプリケーションおよび他のコンポーネントはポート番号を分けてください。
- ・SAS 9.2のアプリケーションブローカとロードマネージャで、旧バージョンのSASとSAS 9.2を制御することを推奨します。

その他、1つのbroker.cfgファイルに旧バージョンのSASとSAS 9.2サービスの定義を含めることもできます。なお、1台のマシンに旧バージョンのSASとSAS 9.2のCGIツールをインストールすることが可能です。ただし、それらは必ず別のディレクトリにインストールしてください。

##### SAS/ACCESS Interface to PC File Formats

SAS 9.2をインストールするとMicrosoft Data Access Components（以下MDAC）がアップグレードされる場合があります。SAS 8.2にてSAS/Access to PC File Formatsをご利用の場合は、SAS 9.2をインストールする前に、修正ファイル(hotfix)82PF05のインストールが必須となります。SAS 9.2のインストールの前に、SAS 8.2のHot Fix 82PF05を適用してください。Hot Fix 82PF05は、下記のURLよりダウンロードできます。

[http://ftp.sas.com/techsup/download/hotfix/82\\_win\\_dbcs.html#82PF05](http://ftp.sas.com/techsup/download/hotfix/82_win_dbcs.html#82PF05)

現在適用されているHot Fixの番号は、下記のフォルダに保存されている.audファイルの名前で確認できます。

```
!SASROOT¥core¥sasinst¥hotfix
```

（!SASROOTはSAS 8.2がインストールされているフォルダを表します。）

.audファイルのファイル名は「Hot Fix番号wn.aud」のようになっています。

例 82pf05wn.aud

下記ページにSAS 8.2とSAS 9の共存に関する情報があります。

<http://support.sas.com/rnd/migration/papers/peaceful.html>

## 5. 移行の環境準備および移行作業

SAS 9.2へのインストールが完了した後、移行作業を開始します。移行作業にあたり、SASシステムオプションの標準値を変更し、SASを使用している場合がありますので、オプションを旧バージョンで変更している場合、SAS 9.2の移行環境へ反映させてください。

代表的なオプションとしては、以下の通りです。

- ・ YEARCUTOFF /\*Y2K 対応 /\*
- ・ MEMSIZE /\*SAS使用最大メモリ容量 /\*
- ・ BUFNO /\*ファイル/Oの最大バッファ数 /\*
- ・ BUFSIZE /\*ファイル/Oの最大バッファサイズ /\*
- ・ WORK /\*SAS一時ファイルの作成先 /\*

また、アプリケーションで使用する入出力ファイルの割り当てを自動化するため、自動実行ファイル(AUTOEXEC)で設定している場合があります。

### 5.1 SASシステムオプション

SASシステムオプションの設定は、環境設定ファイルで行いません。また、この環境設定ファイルは、SAS/CONNECTスクリプトファイルや、SAS起動コマンドでもSASシステムオプションを動的に指定することもできます。SASインストール時に作成される環境設定ファイルはもちろんのこと、SAS起動コマンド中に、「-CONFIG ファイル名」の指定がされている場合、該当のファイル中に設定されているSASシステムオプションをご確認ください。

##### SASシステムオプションの確認

旧バージョンにて設定されているSASシステムオプションの内容は、OPTIONSプロシジャで確認できます。

```
PROC OPTIONS; RUN;
```



### 5.2 自動実行ファイル(AUTOEXEC)

自動実行ファイルが使用されているかを判断するには、環境設定ファイル中に「-AUTOEXEC ファイル名」の設定がされているか確認します。この指定がされている場合、自動実行ファイルを使用しています。また、この自動実行機能の設定方法は、環境設定ファイルと同様に、SAS起動時にも有効ですので、SASの起動コマンドやSAS/CONNECTスクリプトファイルでも「-AUTOEXEC ファイル名」を使用しているか、合わせてご確認ください。

### 5.3 SASライブラリ(データセット、ビュー、カタログ)の移行

旧バージョンのSASライブラリを移行するには、基本的にMIGRATEプロシジャで全て移行できます。ただし、特定の条件によっては、CPORT/CIMPORTプロシジャを使用する場合があります。詳細は、次号Technical Newsの特集でご説明させていただきます。

### 5.4 その他の注意点

ストアドマクロの再コンパイル

ストアドマクロを格納しているカタログは、移行することができません。SAS 9.2にてストアドマクロを使用する場合は、ソースコードから再コンパイルを行ってください。

SASコードに関する注意点

SAS 9.2で、既存SASプログラムの移行テストを実施する前の事前確認といたしまして、既存のSASバージョンでプログラムの実行時にワーニングメッセージが出力されている場合、このメッセージが出力されないよう、適宜SASプログラムの修正を行ってください。

### 5.5 新環境への切り替え

SAS 9.2のテスト環境で、既存のSASアプリケーションを実行し検証した後、本番環境へ切り替えるための最終稼働確認を行いません。確認事項は、以下の通りです。

- ・ SASアプリケーションが本番環境で正常に稼働する
- ・ SASアプリケーションの実行結果が移行ゴールに合致している

特に、移行テスト環境と本番環境にシステム上などの差異がある場合や、移行テスト環境や本番環境で使用したデータ内容や容量に大きな差異がある場合は、注意が必要です。

その他、SASアプリケーションの運用に関連したドキュメントや、設計資料などを作成されている場合、必要に応じて適宜修正および更新などをご検討ください。

## 6. おわりに

本稿では、SAS 9.2( Foundation )への移行の概要、方法論、注意点についてご紹介しました。お客様がSAS 9.2への移行を検討する際などの参考にしていただければ幸いです。

# SAS Academic News



## はじめに

大好評のユーザー事例同志社大学宿久研究室も三回目の掲載となり、今回で最後の掲載となりました。一年間ありがとうございました。四方山話と社会学系におけるSASは引き続き専門的な内容でお届けします。

事例紹介

コラム「SAS四方山話」  
カスタマーリレーション

## 事例紹介

宿久 洋

同志社大学文化情報部/文化情報研究科  
教授 博士(工学)



SAS: 学生様は何名くらいいらっしゃいますか

宿久: ゼミ生が10名、学部の学生は250名(来年から280名になります)、新カリキュラムが走って、データ分析科目で非常に充実させています。そのデータ分析のところと数理的なところの科目とをあわせて運営される予定です。その中にソフトウェア、統計ソフトウェアを含めた演習が組み込まれていくんですが。

SAS: 統計ソフトというのはSASだけではないのですか?

宿久: SASだけでなくどのソフトも導入し使用します。1年の場合はだいたいエクセルをベースに行なうし、演習系のものでSPSSを使用します。一番高度な分析的なものにSASを使用します。ですから、一年生ぐらいですとSASのEnterprise GuideとSPSSがセットで運用されますが、基本的には人文系のひとたちのシェア・経済系のシェアは主にSPSSを使用しています。

SAS: 卒業論文等に最も多く使われるソフトは何ですか?

宿久: それは指導の先生方に依存します。先生方が育ててきたバックグラウンドで左右されるようです。例えばSPSS、SAS、Rなど多種多様です。それは先生あるいはその指導員の学生さんの話し合いで決めます。わたしのところは、SASまたはRでしょうかね。Rがその用途で得意とするところはRで使います。そうですね、最近ではSASとRを使うのはわれわれのトレンドですよ。

SAS: 最近Rはよく耳にしますよね。

宿久: SASがまだ入っていないところはRを使っているんじゃないでしょうか。細かいところもRでできてしまいますよ。いろいろですね。

SAS: 宿久ゼミでもEnterprise Guideはお使いになっていますか。

宿久: 僕のところは使うというよりは、学生さんに「こんなものありますよ」と案内するくらいです。コースや演習の中で紹介をして、あとは学生がそれを使用して分析をするかしないかでしょうか。基本はSAS言語を書いて行っているようです。Enterprise Guideを前面に出して「これでなにかやれ」ということではないですね。いろいろな考え方がありますが、Enterprise Guideというアプリケーションができたおかげで他のソフトを使わずにSASを通すこともできるようになりました。今までは通すことすらできなかったんですよ。そういう意味ではSASでの選択肢が増えたってことでしょうかね。あと、SASの本も何冊か出てきましたね。ゼミ室にも置いてありますよ。学生がそれを読んで参考しているようです。やっぱり本ですよ。Rですけど、課題に特化した本がたくさん出ています。ここにも何冊も置いてあります。Rの本がだいぶ増えてきたので、特にバイオインフォマティクのように特化したものがあると、その路線の方々はそれだけで手に取りますよね。そのほかデータマイニング関係の本もありますよ。SASは英語の本はたくさんあるんですが、やはり日本語化されていないのが痛いところですね。ただ、SASのテキストはとても良いです。講習会のテキスト(SAS社で行っているSASTレーニング)は非常に良くできています。あれを大学に小売していただくと助かりますがね。講習会のテキストがみなさんの手に届くようになれば本は必要がないと思います。

SAS: 米国のテキストは頻繁に使っていますか。

宿久: Trainer's Kitですよ、あのお陰でゼミや授業は助かっています。英語ですが、それでもとても便利です。SAS社としては、Trainer's Kitは売り物なので取り扱いが難しいのでしょうかね。Academic Programを導入した大学への特権でもありますよね。

SAS: はい、Academic Programを契約いただいた先生方にしかお出ししていません。

宿久: 逆に、SAS講義を普及させるためにTrainer's Kitを盛り込んだ個別のライセンス体系も考えていただけたら、SASが導入できないほかの大学にも普及するんじゃないでしょうかね。Trainer's Kitですが、あれを先生方が自分で作るとなると敷居が高いんですよ。先生方も多忙ですからあの内容のものを作るとなるとかなり時間が掛かります。あれはコースを受けてみてもわかるのですが、とても考えて作られたものだと思いますよ。

SAS: そうですね、サンプルデータも揃っているし、演習の体系もよいバランスがあると思います。本社の狙いとしては、アカデミック

にはSASの授業を開設頂く先生方への支援を強化する方針になっています。

宿久: ええ、西海岸の講習を受けたときもそれは感じました。教材の無償配布やテキスト教材の支援など、こんなにサービスしてくれていいのか(笑)って思いましたよ。だからいろいろな国から来たいという人が多いんですね。ヨーロッパやアジア圏からも来ていました。

SAS: 本社の指示のもと、日本語版ではお出ししていないのですが英語版ではお渡ししています。われわれとしてもSASの講義が増えただけであればありがたいと思っています。将来的にはその講義を受けた学生さんが社会に出てSASを使っていたら何よりです。

宿久: SASは大学で使える価格になってきましたね。でもわたしが思うにまだまだ学生にSASは驚沢な感じがします。

SAS: Academic Programは他のソフトの会社の価格体系と比較しても安価に設定しています。そのため徐々にですが普及してきました。ただ、SASを使ってもらわないといくら安くても導入した意味がないので解約する大学もあります。

宿久: いろいろな考えがありますが、大学によってまちまちなんですよ。おらかな大学もありますし、今ですと予算削減が大きな課題である大学が多いでしょうからね。ですが、SASのAcademic Programを導入する大学が増えているということは競争に勝っているってことなんじゃないでしょうか。

SAS: そうだとありがたいのですが、こちらの大学様の1年生、2年生でSPSSを使うという話がありましたが、SASの狙いは、そのあたりからSASを何とか使用してもらえないか、というのが課題です。

宿久: (その課題は)あるでしょうね。ですが、SASをそのくらいの学生に使用させるといった場合、何かあと一押しが必要でしょうね。それは民間企業への普及もそうでしょうし、人文系の先生方に、必ずデータ解析をするという認知が、なぜかSPSSでするといった流れになっているようです。その人たちが学んだときのソフトがSPSSなんですね、それがいま学生を教える立場になって、SPSSを教えているという流れですね。ですから、アカデミックは怖いんですね。あるとき気づいたときにはバツと情勢が変わるといったところが恐ろしいです。

SAS: SAS Japanとしては、本社からの最新情報を大学の先生方へ提供したいと考えています。あとは、今回もEnterprise Guideのマニュアル本が出ましたが、こういった本をシリーズ化して何冊か発行していく予定です。それと、民間企業でSASを使っている団体と大学間の産学連携に貢献したいと思っています。

宿久: そうですね、せっかく大学でSASを学んだがその用途が明確でないという学生へのやる気にも影響があると思います。みんなが使えるといった理想の話はしないけれども多少なりとも就職していく学生へのメリットがあるとよいと思います。間にSAS社が入っていただけると助かりますね。日本はあまり見受けられませんが海外ではそのあたりは力を入れているようです。問題を持ち込まれてある程度アドバイスするとか、そのほうがうまくいくことがありますね。あるいは学生を送り出した後など、というのが形になっていくとよりよいのかと思います。そうすると他の会社にはないメリットがありますね。

SAS: そうですね、バリューみたいなものをご提示できるかと思います。

宿久: SASの強みは大手企業に導入しているところだと思います。学生さんもそのようなところに目をつけていると思います。そこをうまくSAS社にやっていただけたらよいと思います。

SAS: 実際、企業でどのように使われているか、などご興味はございますか?

宿久: もちろんあります。前回も同じ京都の大手通販会社に大学に来ていただいて社内におけるSASの必要性をお話しいただきました。キャリアセミナーみたいな形で続けてしまえば、あまりSASを強調することなく学生にアピールできると思います。うちの学科は、SASを前面に出してもよいんですけどね。SASに力を入れることは、データサイエンスを専攻している学生のモチベーション維持に非常に役に立ちます。

SAS: お客様にお話を聞くとSASのエンジニアがいない、統計がわかる人がいない、SASを使える人がいない、といった課題があります。そういう企業側が抱えている問題もかなりあります。そういった課題が大学の先生方と一緒に連携していただいて、就職斡旋ということまでにはなりません、企業、大学の連携を助けたいと思います。学生さんもSASを学ぶ意識が変わりますよね。

宿久: われわれはSASを含めたデータ分析が特殊な行為ではなくて非常に日常的に必要なことだと思っています。というようなことを義務付けさせたいと考えています。文系の学生にも、今後は企業に出て役に立つということを言っていけないといけないと思います。うちの学部は文系の統計学部といった感じのカリキュラムで作られています。実際データ解析するのは理系出身の方々ではないんですよ。むしろ文系のほうが社会に出てデータを扱う職種に付くことが多いわけなんです。データ分析力が今後、何系出身に関わらず必要になる時代ではないかと思えます。SASに要望するのは、企業のたとえばマーケティング部でどのように使われているかを大学側にフィードバックがほしいですね。

SAS: そうですね、企業側では情報活用がキーワードになってきています。データサイエンスというところに注力をあてて、せっかく集まった情報を生かそうという動きにどんどんなっています。将来の予測、そこに重きをおく企業が増えてきています。

宿久: 意思決定の道具として、勤や経験ももちろん大事ですけれども、このような定量的なものを分析する時代がやっきたのか、と思えますね。いわゆる教育の分野での例示されているのは、たとえばヒストグラムの読み方がわからない、棒グラフの読み方がわからない、というのを聞きます。いわゆる計算力でなく情報の処理解釈が足りない、ということで指導要領が今度変わるのだから教育のシステムを統計とかデータ処理に力を入れる時代になります。中学校で、いわゆる「資料の整理」というのが入りますし、「資料の活用」という項目があるんです。それが「資料の分析」として入ってくる。もしかしたら必修の中に入ってくるかもしれない。

SAS: 中学・高校で情報処理の解釈の基礎ができれば、SASは学部生にも問題なく吸収できるソフトになりますね。

宿久: 使う側からすると、ひとつのソフトウェアで何をやるっていう感じではないんですよ。SASで処理をするところ、エクセルで処理をするところ、または他のソフトで処理するのが早い場合はそれを使います。この部分はこれ、この部分はこれ、といった感じです。こだわりなくやっています。どちらにしても、情報の処理を中学・高校でする場合、教える先生方が変わっていくといけないと思います。

果たして先生方がそのような環境についていけるか、というのが大きな課題になるんじゃないでしょうか。数学の先生などは、理論は教えられるが、データに触ったことがない、なんて話はよくあります。

SAS: 学生の就職先についてお聞きします。たとえばどのようなところに就職を希望されている学生が多いのでしょうか。

宿久: 製薬会社が何名か、金融が何名か、でしょうか。製薬系の統計解析部などへ希望しています。金融はなかなか難しいようです。中には頑張って就職するものもいます。あとはシンクタンク系でしょうか。教える側として、企業での要求が大きい、かつ能力を要求される業界はあまりお勧めしていませんが、しかしやりがいはあるというところであれば学生が希望すれば応援します。その他、保険業界でしょうか。損保・生保ですね。これはアクチュアリを取っていききたいというところですね。

SAS: 保険業界も、ソルベンシーIIという規制が入る予定です。その規制が入るとリスク管理などの分析が必要になります。

宿久: もともとアクチュアリをおかなくてはいけないことになっているので、そういうところをターゲットにしています。

SAS: いわゆる金融機関ですと、パーゼルという規制がありますね。それでオペレーショナルリスク管理と信用リスク管理、市場リスク管理、この3つの管理をしなくてはならないのですが、市場リスク以外の2つというのは、完全に統計です。お陰様でSASがそこでデファクトになっているため、メガバンク・地銀などSASを導入いただいている経緯があります。

宿久: そのような話は、学生さんに伝えてもらうとSASを学ぶモチベーションに繋がりますね。Technical Newsも学生で購読しているのは少ないようです。

SAS: 今ですと弊社ホームページの教育機関のサイトで最新情報を公開しています。わたくしどもも、本社には有用な情報があるにもかかわらず、日本サイドでそれを伝え切れず、今後教育機関のホームページで情報公開に努めていく予定です。もしくは、地域に限定したユーザー会やコミュニティーの設立も考えています。

宿久: 東京で毎年夏に行われているユーザー総会は数居が高いんですよ。ですが、あれはあれで他のメーカーさんの目標になってしまっていて、他ベンダーがSASにまねてユーザー主催の会を開いていますよ。

SAS: そうですね。わたしたちとしては東京に限らず、関西エリアを限定したユーザー会も企画し交流の場を作りたいと思っています。

宿久: そうですね、必ずしも成果報告というものでなくても、企業と教育の場のものが交流を持つ場があるとよいですね。こういったのは、文教の先生方が興味をもたれる方が多いと思います。

SAS: 学生が論文発表をする場も作る予定です。

宿久: ぜひ開催していただきたいですね。

SAS: ありがとうございました。

## コラム「SAS四方山話」

### 第4回 「アンケート四方山話(2) 実施～集計まで」

大橋 渉

東京医科歯科大学 歯学総合研究科大学院 生命情報学  
情報医科学センター 臨床・生物統計グループ 特任助教

#### 1) 前回は振り返って

前回までにアンケート(調査票)の作成方法についてお話させていただきました。読者の方からも、「言われてみれば訳の分からないアンケートに回答した覚えがある。自分たちもそこまで気を付けて、訳の分からないアンケートを作成しないようにしたい」などのご意見を頂けたのは嬉しい限りです。前回も書きましたが、やはりSASユーザーの皆様は熱心な方が多いのだと改めて感じる次第です。今回は「アンケートの実施から集計までの流れ」についてお話させていただきます。結果は集まったものの、世の中には処理方法が分からないまま眠っているデータもまた多いようです。では、データの賞味期限\* )が切れてしまう前に取り掛かりましょう。

\* )例えば政府系の調査などは、調査時点から2～3年も経過した後に結果が出てくることが多いです。大規模調査ならば止むを得ないことが多いですが、可能であれば早めに処理してしまうことをお勧めします。集計・解析が終了したときには、「もう結果が不要になっていた」「時代が変わってしまっていた(?)」などということの無いように…。

#### 2) 調査票配布後の仕事は?

アンケートの方法によっても異なりますが、配布から集計までには結構やるべきことがあります。例えば郵送調査の場合ですが、配布後数日間で何らかの反応があります。多くの調査では「問い合わせ先」の連絡電話番号が記載されておりますので、まずは回答者からの問い合わせに奔走することになります。続いて回答拒否や居所不明により戻ってきた調査票のチェックがあります。拒否及び不明の場合は、次回以降の調査において対象から除外することになります( 郵送料も助かります )。特に拒否の場合には その原因を調べる必要があります( うっかり対象者に電話などをすると「感情的に」返答されることもありますので要注意です)。開封もされずに「面倒くさい」という程度の理由であれば、回答者個人の問題として処理できる程度ですが、開封後に「面倒くさい」という理由がやたらと多い場合、そのときには 調査票の構造自体の問題を疑いましょう( 前号を参照 )。可能であれば、日別に有効回答数、無効数とその内訳を集計して、余りにも回収数が少なければ、督促などを検討することをお勧めします。督促の方法は、昔は電話やハガキによる督促が主体でしたが、近年では電子メールも用いられます。督促による回収率への貢献はあることはありますが、せいぜい数%程度と考えておきましょう。また筆者の経験では、督促後の回答には「どちらでもない」がやたら多く、記述部分は無回答率が高いなどの傾向が強かったことを記憶しておりますので、場合によっては回答の精度自体が低下することさえあります\* )。という訳で、督促の実施は督促方法や予算、工数などと相談して検討することをお勧めします。

以上、調査の締め切りまで上記の ~ を繰り返すわけですが、そのうち の調査票の構造についての検討結果は次回以降の調査に役立てるようにしましょう。調査の最中に項目を変更してしまうと、それ以前の結果との整合性が取れなくなってしまうといった問題が発生しますので、もしも変更するのであればその部分は覚悟して行ないましょう。締め切りがやってきたら、最終的な配布数、有効回答数、回答拒否数、不明数などを集計し、各項目について

配布数に占める割合をレポートします。このとき、配布数に占める有効回答数のことを「回収率」と呼び、この割合が高いほど優秀な調査であるとされる傾向があるようです。しかしながら謝礼の有無や調査方法、時期などによって異なりますので、一概に優秀な調査であるとは言えないでしょう。優秀な調査とは、可能な限り短い日数と少ない予算で、研究者の問題意識を解決できた調査に他ならないと思うのですが…?

\* )現在、筆者は過去における社会調査を検討することで「督促と回答内容」の関係について学術的に研究しております。結果はまた報告させていただければ幸いです。

#### 3) 入れ物と中身

さて、ようやく回答結果も集まりました。Web調査などで自動的に集計されているような場合もあると思われませんが、ここでは一般的な郵送・回収などの紙による調査票を、どのように集計するかについて説明させていただきたいと思えます。そもそも集計のためには、収集したデータを集計用にとまとめる必要がありますが、多分多くの皆様がそのためのツールとしてMicrosoft Excelを思い浮かべたのではないのでしょうか? 確かに筆者のお勧めもExcelなのですが、ご注意ください。Excelは「入れ物と中身」のところまでの利用ということ。基本的にExcelは「簡易表計算ソフト」であり、あくまで統計ソフトウェアではありませんので、集計はともかく解析結果には若干の不安もあります。(多くのSASユーザーの方がそのようにされていると思われませんが)ここではSASによる集計・解析用のデータをExcelによって作成する方法について説明させていただきます。まずは下記のような調査項目を想定してみましよう。

Q1:年齢 ( ) Q2:性別 男・女  
Q3:過去に統計ソフトウェアの利用経験がある? はい・いいえ  
Q4:居住地域 1.北海道・・・(中略)・・・8.九州・沖縄  
Q5:この統計ソフトウェアの満足度は?  
1.非常に高い 2.高い方 3.低い方 4.非常に低い  
Q6:統計ソフトウェアを購入するときに何を重視するか?(該当するもの全て)  
1.価格 2.機能の多さ 3.使いやすさ 4.ユーザー数  
5.その他( )  
Q7:統計ソフトウェア全般に対するご意見  
( )

これら7項目の回答を入れる枠組みが「入れ物」であり、入れられるデータが「中身」になります。まずは入れ物のデザインですが、SASは基本的に「タテ方向への演算」を行なうものですので、調査項目をヨコ方向に、回答者はタテ方向に入力していきましょう。まずExcelの1行目には、設問項目(変数と呼びます)を入力するのですが、このとき、SASへの読み込みを考慮するならば「日本語は用いずに」行ないましょう。日本語で入力すると、読み込み時に変数名が左から勝手に「\_COL1,...\_COL」とされてしまいますので、注意が必要です。

次に、設問の種類について考えてみましょう。Q1は回答通りに直接年齢を記載していけば良いのですが、このとき「半角数字」で入力されていることをご確認ください。全角数字が半角に混在しているようなデータの事例も結構見受けられますが、その場合プログラムにおいては全く別物と判断されてしまいますので要注意です。Q2～Q5のうちQ2、Q3ですが、ここでは日本語を直接入力するよりも「男性=1、女性=2」「はい=1、いいえ=0」のように入力規則を決定して、入力のための工数を簡略化するようにしましょう。たとえ入力は数値でも、SASフォーマットを用いての日本語表示は可能ですし、入力ミスを防ぐためにも可能な限り簡略化(手抜き?)することを考えましょう。



Q6は他の項目と異なり「複数回答」ですので、当然入れ物の設計にも注意が必要となります。仮にQ6というセルを一つだけ作成した挙句に、「12345(教本の多さ)」などを入力した場合には、非常に集計が苦しくなります...と言うよりはむしろ、このままでは集計が不可能です。この場合には、Q6の1~5の各選択肢について単独で入れ物を作成し、回答のあった部分について印(フラグ)という言い方が一般的ですが、立てるようにしましょう。例えば筆者の場合ですが、「選択があった場合には1を、無い場合には0を入力する」というような方式を採用しており、学生に対する指導時も「お勤めのやり方」として紹介させていただいております。Q6「その他」およびQ7の文字情報にしましては、SASへの読み込み時に途切れないように注意が必要です(詳細はまた次号にて)。

以上、SASによる集計・解析を意識した入れ物と中身について説明させていただきました。特に、Q6のような複数回答の情報については、筆者の経験ではかなりの頻度で同一セルに詰め込まれている事例を見せつけられて(?)おりますので、是非とも皆様はご注意くださいませ。特に、詰め込まれた状態で統計担当者に「何とかしてくれ!」と持ち込むようなことはやめましょ(笑)!

次回こそは前回の予告通り、解析方法を説明したい次第です。

ここは注意が必要です!

NUM	Q1	Q2	Q3	Q4	Q5	Q6_1	Q6_2	Q6_3	Q6_4	Q6_5	Q6_0t	Q7
0001	28	1	0	3	2	1	1	0	1	0		
0002	35	0	1	2	3	0	0	0	0	1	教本の多さ	特になし

全て数値は半角で入れてくださいね!

## カスタマーリレーション

### 社会学データ分析におけるSASの使用(その3)

労働政策研究・研修機構

平田 周一

#### 1. マルチレベル分析

今回は、近年、急速に注目を集めているマルチレベル分析(Multi Level Analysis)を取り上げる。

マルチレベル分析については、ほかにさまざまな呼び名がある。なかでも有名なのは、「階層線形モデル(Hierarchical Linear Model)」というもので、アメリカの教育について研究しているBrykとRaudenbushによる「Hierarchical Linear Models: Applications and Data Analysis Methods(2002, Sage Publication)」によって有名になった。この2人は、HLMというマルチレベル分析専用のソフトウェアを開発している。統計学研究者の間では、しばしば、混合モデル(Mixed Model: 例えば、Little et al. 1996 *SAS system for Mixed Models* SAS Publishing)という言葉が使われている。個人的に気に入っているもので、コンテクスチュアル分析(Contextual Analysis G. R. Iverson 1991 Sage Publication)という言い方もある。

どのような呼び方をされるにせよ、マルチレベル分析とは、分析対象の単位を複数仮定した重回帰分析、あるいは重回帰分析に類似したものである。

分析対象の単位を複数仮定するとはどういうことなのか。簡単にいえば、分析対象の標本があるグループに属する場合、個々の標本を一つのレベル、個々の標本が属するグループを一段階上のグループと考えるということだ。上川一秋氏による簡潔な解説(上川一秋 2006「異なる分析レベルの因果を同時に考える:階層線形モデル(HLM)」)と謝野有紀他編『社会の見方、測り方』勁草書房所収)によれば、マルチレベル分析は、古典的な重回帰分析と基本的には変わらない。ただし、従来の重回帰分析は標本間の独立性を仮定していたが、「社会学で用いられるデータにおいて、この仮定が成り立たないことが多くある(前掲書、p.122)」。例えば、学校ごとに生徒を選んで、個々の生徒を対象に調査を行なう場合、同じ学校に通う生徒は何らかの特徴を共有していると考えられるので、他の学校の生徒と比べて、よく似た傾向が現れる。このような場合に用いられるのがマルチレベル分析である。マルチレベル分析では、個々の標本における従属変数(例えば生徒の成績)と独立変数(例えば親の社会的地位)の間の回帰係数および切片の値が、より上のレベル(例えば学校)によって異なると仮定し、上のレベルの変数によって回帰係数や切片の値がどのように異なるかを分析することができる。このようにして、学校ごとの教育方針の違いによる影響を考慮することができる。また、国際比較調査などでは、国ごとの制度や文化の違いを考慮した分析を行なうことができる。制度や文化、すなわち社会的コンテキストが異なることを考慮して分析を行なうことができることが、コンテクスチュアル分析と呼ばれる所以だ。

マルチレベル分析では、どのようなモデルを立てるのかを簡単に説明しよう。通常の重回帰分析モデルは、次のようになる。

$$Y_i = \alpha + \beta X + e_i$$

ここで、 $\alpha$  は定数項、 $\beta$  は独立変数 $X$ の効果、 $e_i$ は誤差項を示す。誤差項と従属変数に添字が付いているのは、 $i$ 番目の対象の誤差項によって従属変数の値が変化するという点を強調するためだ。一方、定数項( $\alpha$ )と独立変数の効果( $\beta$ )の値はどの対象者に対しても共通で変化しない。

マルチレベル分析では、定数項や独立変数の効果が、社会的コンテキストの違いによって変化すると考え、次のようなモデルを立てる。

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_1 + e_{ij}$$

$$\beta_{0j} = \beta_{00} + u_{0j}$$

$$\beta_{1j} = \beta_{10} + u_{1j}$$

定数項を  $\beta_{0j}$ 、独立変数  $X_1$  の効果を  $\beta_{1j}$  に表記を変えている。また、添え字に  $j$  という文字が一つ増えているが、これは、一つ上の階層が個のグループに分かれていることを示している。2番目と3番目の式は、それぞれ、通常の回帰モデルの定数項と独立変数の効果が、それぞれの平均 ( $\beta_{00}$ ,  $\beta_{10}$ ) を中心に分散があることを示している。

さらに分析を進めて、独立変数の効果が一つ上の階層におけるグループが持つ変数によって異なるとするモデルを想定する。

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_1 + e_{ij}$$

$$\beta_{0j} = \beta_{00} + u_{0j}$$

$$\beta_{1j} = \beta_{10} + \beta_{11}X_j + u_{1j}$$

3番目の変数  $X_1$  の効果  $\beta_{1j}$  を説明する式で、新たに  $X_j$  という変数を設定し、それは  $\beta_{10}$  という効果を  $\beta_{1j}$  に与える。最も低い階層を示す最初の式に、さらに独立変数を加えることもできる。

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_1 + \beta_{2j}X_2 + e_{ij}$$

$$\beta_{0j} = \beta_{00} + u_{0j}$$

$$\beta_{1j} = \beta_{10} + \beta_{11}X_j + u_{1j}$$

$$\beta_{2j} = \beta_{20}$$

今度の例では、 $X_2$  という新しい独立変数が最初の式に加えられているが、その効果は最後の式に示されているように、上の階層におけるグループごとでの違いはなく、全ての対象者に対して同じ値を持っていると仮定されている。新しい独立変数  $X_2$  は、通常の回帰分析における独立変数と同じ扱いになる。通常の回帰分析では、定数項と独立変数の効果は、全ての対象者に対して同じ値を持つ。これを特別に固定効果 (fixed-effect) と呼ぶ。一方、マルチレベル分析における定数項と独立変数の効果の値は、それぞれ誤差項 ( $u_{0j}$ ,  $u_{1j}$ ) を持つことが許される。このような場合、これらをランダム効果 (random-effect) と呼ぶ。さらに、最後の例のように、マルチレベル分析では一つのモデルの中に固定効果とランダム効果の両方を設定することができる。これが混合モデルとも言われる所以だ。

一つ上の階層におけるグループとして、さまざまなものが考えられる。上川 (2006) では、親の教育レベルと子供の成績の関係をアメリカと韓国で比較する例が示されている。この場合、国の違いが一つ上の階層におけるグループとなる。国際比較研究でなくとも、地域の違いがどのような影響を及ぼすかを分析することもできる。ルーグ (Douglas A. Luke 2004 *Multilevel Modeling* Sage Publishing) は、議会ではばくち産業に有利な投票をするという行動が、議員の出身州によって大きく異なることを示している。前述の、ブライクとローデンブッシュ (Bryk & Raudenbush 2002) は、親の社会的地位と子供の成績の関係の分析に、学校の種類 (公立であるか私立であるか、あるいはカトリック系の学校であるかなど) による違いを考慮したモデルをたてている。

社会学、あるいは一般的な社会科学で、マルチレベル分析が有効である理由として、調査データの収集方式における根本的な問題があるのではない。

社会科学では、多くの場合、医療研究や理科系の研究のような厳密な実験データをとることができない。実験の代わりに、社会調査を

行なうわけだが、多くの場合、特に全国調査などでは、層化二段抽出法という方法を用いて標本の抽出を行なう。最初に対象者の住む地域を選び、選ばれた地域から無作為に対象者個人を抽出するという方法だ。このようにして選ばれた標本は、居住する地域ごとにある特性を持っていることが考えられる。以前は、このような可能性に目をつぶって分析が行なわれてきた。しかし、マルチレベル分析によって、このような欠点を逆手にとって、プラスの価値を持つ分析が行なえるようになった。地域ごとのマクロデータを活用して、既婚女性の就業行動を、地域の保育所などの施設の充実度を考慮しながら分析できるようになった。マルチレベル分析によって、マクロデータとマイクロデータを結びつけることができるといわれているのは、こうしたことが可能であることに由来する。

## 2 SASを用いてマルチレベル分析を行なう利点

SASを用いてマルチレベル分析を行なう利点はいくつか考えられる。まず、マルチレベル分析は、従来の分析方法に比べて斬新なものであり、前にふれたHLM以外にも、マルチレベル分析専用のソフトが開発されているが、多くの場合、SASで用意されているプロシジャーによって、マルチレベル分析を行なうことができる。前に紹介した上川-秋氏のホームページ (<http://www.estat.us/index.html>) では専用ソフトを使わずともSASでマルチレベル分析に十分対応できるという例と実際のSASの使用方法が紹介されている (アメリカ、Optimal Solution Group 上級研究員上川-秋氏には、ホームページからの引用の許可をいただいた。記して感謝したい)。

マルチレベル分析を行なうために、専用のソフトが作られている。前述したHLM®がよく用いられているようだが、その他にもいくつかの専用ソフトがある。しかし、これらの専用ソフトに引けを取らないプロシジャーがSASには用意されている。通常マルチレベル分析を行なうには、PROC MIXEDを用いなければならない。私が、幾つかのマルチレベル分析の教科書に載っている例を、このプロシジャーを使って再現を試みたことがあるが、掲載されている例とほぼ同じ結果を得ることができた (マルチレベル分析のアルゴリズムは複雑なので、ソフトによって、結果の数字に若干の違いが生じることがある)。前述の上川氏のホームページには、マルチレベル分析を行なうためのプロシジャーとして、PROC MIXED以外に、PROC GLIMMIX、PROC NLMMIXEDを活用する例が紹介されている。前者は、一般化線形混合モデル (Generalized Linear Mixed Model) のために新たに設けられたプロシジャー、後者は非線形混合モデル (Non-Linear Mixed Model) のためのプロシジャーで、従属変数がロジット、多項ロジットなどの時にはこれを用いる。

また、マルチレベル分析では、前回紹介したPerson-periodデータを用いなければならないことがあるのだが、前回紹介したように、SASを使えばPerson-periodデータを簡単に作成できる。Person-periodデータを用いたマルチレベル分析は、イベントヒストリー分析のように、時系列で変化する事象の分析に用いられる。以前、引用したシンガーとウィレット (Judith D. Singer & John B. Willet 2003 *Applied Longitudinal Data Analysis Oxford*) では、時系列で変化する事象の分析方法を示したもののだが、前半はマルチレベル分析、後半はイベントヒストリー分析が紹介されている。

時系列で変化する事象の分析でマルチレベル分析を用いる場合、一つ上の階層とされるのは時系列上の時点だ。シンガーとウィレット (Singer & Willet 2003) では11歳から15歳までの間での逸脱行動 (試験のカンニング、薬物の服用、アルコールの飲酒、盗み等) の発生についての分析を例としている。前回紹介した、イベントヒストリー分析との違いは、イベントヒストリー分析がある事象が発生したか否かを従属変数としているのに対して、この例では、時点ごとに行なわれた逸脱行動の数が従属変数となっている点だ。

このような分析でマルチレベル分析を用いなければならない理由は、年齢が上がると逸脱行動が増えたり、減ったりしている。また、最初は低く、中間時点で高くなり、その後また低くなるというように。対象者によって年齢と逸脱行動の関係が多様すぎる点にある。また、年齢時点ごとに独立変数(親との同居、友人関係など)が調べられているが、これらの変数と逸脱行動の関係も年齢によって異なっている。そこで、時点を一つ上の階層とし、時点ごとに共通の独立変数(親や友人についての変数)と従属変数の関係を吟味する、と同時に、時点と逸脱行動の関係を吟味することができるマルチレベル分析が必要となる。

前にも紹介したが、同書の著者は、UCLA内Webサイトで、同書の中で示されている例をさまざまなソフトウェアを用いて再現する例を紹介している(<http://www.ats.ucla.edu/stat/sas/examples/alda/default.htm>)。これを見ると、マルチレベル分析については、HLMなどのマルチレベル分析専用ソフトと同等の機能をSASが持っていることがよくわかるだろう。

### 3 おわりに

三回にわたって、社会学を専攻する筆者がどのようにSASを使っているか、また、使おうとしているかについて述べてきた。SASの持っている利点として、多機能であること、柔軟なプログラミング能力を備えていることなどを述べたかったのだが、十分、筆者の意図は伝わっただろうか。

SASは多機能であり、前回紹介したイベントヒストリー分析、今回紹介したマルチレベル分析を行なうに当たっても十分な機能をもっている。特に、今回紹介したマルチレベル分析に関するプロシジャーの整備は、筆者の知る限り、総合的なソフトウェアの中で、SASにおいて最も迅速に行なわれている。また、これらの分析のために必要なPerson-periodデータの作成には、SASの柔軟なプログラミング機能が大いに役に立つ。

連載の中で度々紹介したが、海外、特に米国の統計に関するWebサイトで、非常に多くのSASに関する資料を見ることができる。少なくとも米国では、社会科学においてもSASが広範に使用されていることの証だろう。この連載の隠れたテーマは、「インターネットでSASのリソースを見つけよう」ということだったのかもしれない。インターネット上に限らず、SAS Publishingからのものをはじめとして、海外では、多くのSASに関する本が出版されている。これらが、日本でも翻訳されればSASに対する認知度も変化するのではないだろうか。

# Q&A



CSVファイルに出力する際、ラベルを出力したい  
外部ファイルの任意の行からデータを読み込む方法  
改行コードについて

Enterprise Guideから外部ファイルを操作する方法

YYYYMMDDhhmmss/YYYYMMDDhhmmssfff形式の日時データを読み込む方法

SAS 6にて圧縮したデータセットを圧縮した状態で移行する方法

Accessデータベースのテーブルに主キーを設定したい

SAS Enterprise Miner 5でバギングやブースティングを実現する方法

モデルパッケージのリポジトリからの削除方法

ROC曲線の比較について

行列とベクトル間の演算

データセットの一部をランダムに表示する方法

効果の名称が長い場合、すべて表示されない

## Q

EXPORTプロシジャーを使用して、データセットをCSVファイルに出力しています。1行目に変数名ではなく、ラベルを出力することはできますか。

## A

EXPORTプロシジャーでは、ラベルを出力することはできません。しかしながら、PRINTプロシジャーでLABELオプションを使用すると、ラベル表示が行なわれます。ODS CSVステートメントと併用することでご要望の処理が可能です。

例

```
DATA x;
  LABEL x='aaa';
  LABEL y='bbb';
  INPUT x y;
CARDS;
1 2
2 3
3 4
;
RUN;

ODS LISTING CLOSE;
ODS CSV FILE="c:\test.csv";

PROC PRINT DATA=x LABEL NOOBS;
RUN;

ODS CSV CLOSE;
ODS LISTING;
```

## Q

1行目にヘッダー情報があるCSVファイルを、データ部分のある2行目から読み込みたいのですが、何か良い方法がありますか？

## A

INFILEステートメントのFIRSTOBSオプションで、外部ファイルの読み込み開始行を指定することができます。DATAステップのSETステートメントに、データセットオプションとして指定できることと同様、外部ファイルの読み込みに使用するINFILEステートメントにも指定できます。

例

```
INFILE "C:\sample.csv" FIRSTOBS = 2;
```

**Q** EXPORTプロシジャ等で、SASデータセットを出力する際、Windowsの改行コード[CR+LF]ではなく、UNIXの改行コード[LF]に変換して出力することができますか？

**A** TERMSTRオプションにLFを指定し、FILENAMEステートメントを定義することで可能となります。

例

```
FILENAME outfl 'c:¥temp¥test.csv' TERMSTR=LF;

PROC EXPORT DATA=sashelp.class
  OUTFILE = outfl
  DBMS = csv;
RUN;
```

**Q** SAS FoundationからはXステートメントとオペレーティングシステムのコマンドで外部ファイルを削除していますが、Enterprise Guideからは、オペレーティングシステムのコマンドが利用できません。何か良い方法はありますか？

**A** 通常、Enterprise Guideでは、外部コマンドの実行が制限されており、ファイルの削除であれば、FDELETE関数の利用をご検討ください。

例

```
/* ファイル X:¥test.xls を削除する例 */
FILENAME target "X:¥test.xls" ;
DATA _NULL_ ;
  RC = FDELETE("target");
  MSG = SYMSG();
  IF RC ^= 0 THEN PUT RC= MSG=;
RUN;
```

**Q** YYYYMMDDhhmmss/YYYYMMDDhhmmssfff形式の日時データについて、読み込む方法を教えてください。

**A** ND8601DTインフォーマットを指定することで、読み込みが可能です。次の例では、ND8601DTインフォーマットによる読み込みと、DATEPART関数によるSAS日付値の生成を行いません。

例

```
DATA sample;
  INPUT dt:ND8601DT.;
  date=DATEPART(dt);
  FORMAT dt DATETIME22.3
         date YMMDD10.;

CARDS;
20080416201953700
20080417211953600
20080229231953999
20080301101953
;
RUN;
```

**Q** SAS 6にて圧縮したデータセットをSAS 9.1に移行したところ、データセットの容量が、急激に増加しました。以前と同じように圧縮した状態で保存したいのですが、なにか良い方法はありますか？

**A** SAS 6にて作成された移送ファイルは、データセットの圧縮属性を保持しません。こちらはSAS 6で作成された移送ファイルの仕様になります。次のSAS Noteに詳細がございます。

<http://support.sas.com/kb/4/720.html>

SAS 6にて作成された移送ファイルを圧縮された状態でCIMPORTするには、COMPRESSシステムオプション、またはLIBNAMEステートメントのCOMPRESSオプションを使用する方法がございます。以下の例をご確認ください。

例

```
/* COMPRESSシステムオプションを使用する場合 */
OPTIONS COMPRESS=YES;
LIBNAME target "c:¥sas9target";
FILENAME trans V9 "c:¥sas6source¥sas6trans.cpt";
PROC CIMPORT INFILE=trans LIB=target MEMTYPE=data;
RUN;

/* LIBNAMEステートメントのCOMPRESSオプションを使用する場合 */
LIBNAME target V9 "c:¥sas9target" COMPRESS=yes;
FILENAME trans "c:¥sas6source¥sas6trans.cpt";
PROC CIMPORT INFILE=trans LIB=target;
RUN;
```

COMPRESSシステムオプションに関しては全てのSASデータセット、LIBNAMEステートメントのCOMPRESSオプションに関しては定義したライブラリ配下の全てのデータセットを圧縮しますので、ご注意ください。詳細につきましてはSASヘルプなどから、COMPRESSシステムオプション、LIBNAMEステートメントのCOMPRESSオプションをご確認ください。

**Q** LIBNAME ODBCエンジンで接続したAccessデータベース上のテーブルに主キーを設定するにはどうしたらよいですか。

**A** 通常、SASデータセットに対して主キーを設定する場合はDATASETSプロシジャでIC CREATEステートメントを使用しますが、ODBCエンジンライブラリ上のテーブルに対してはIC CREATEステートメントが使用できません。ODBCエンジンライブラリ上のテーブルに主キーを設定するには、SQLプロシジャのEXECUTEステートメントに、主キー作成用のALTER TABLE句を記述する必要があります。以下にサンプルを示します。

例

```
/* データソース testdb を a という名前のライブラリに設定 */
LIBNAME a ODBC DSN=testdb;

/* サンプルテーブルの作成 */
DATA a.test;
  SET sashelp.class;
RUN;

/* SQL プロシジャで pk_name という名前の主キーを name 列に付与 */
```



```
PROC SQL;
CONNECT TO ODBC(DSN=testdb);
EXECUTE (ALTER TABLE test ADD CONSTRAINT
pk_name PRIMARY KEY (name)) BY ODBC;

DISCONNECT FROM ODBC;

QUIT;

LIBNAME a CLEAR; /* ライブラリの解除 */
```

**Q** SAS Enterprise Miner (EM) 4.1や4.3では、グループ処理ノードを使用してバギングやブースティングを行なうことができましたが、EM5.2 やEM5.3ではどのようにして実現可能でしょうか？

**A** EM5.3では、開始グループノードと終了グループノードを使用します。処理を反復するノード群を前記の2つのノードで挟みます。また、開始グループノードのモードプロパティでは、目的に応じてバギングやブースティングを選択してください。なお、EM5.2ではこの機能はサポートされていません。

**Q** SAS Enterprise Miner 5.3において、モデルパッケージを作成してリポジトリに登録しましたが、削除する方法がわかりません。どのようにしたらよいでしょうか？

**A** モデルパッケージマネージャから削除可能です。ただしリポジトリに関しては、サブミットボタンを押すまで画面に現れません。もし正しく削除できない場合には、テクニカルサポートまでご連絡ください。

**Q** LOGISTICプロシジャが出力するROC曲線下の面積 (AUC) を、複数のモデル間で比較する検定を行なうにはどうしたらよいでしょうか？

**A** SAS 9.2からは、ROCステートメントやROCCONTRASTステートメントを使用して、ノンパラメトリックな手法に基づく検定を行なうことができます。詳細については、LOGISTICプロシジャのドキュメントを参照してください。SAS 9.1.3またはそれ以前のリリースでは、米国SAS Institute Inc.のWebサイトで公開されているSASマクロを使用することができます。

Sample 25017: Nonparametric comparison of areas under correlated ROC curves  
<http://support.sas.com/kb/25/017.html>

**Q** SAS 9.1.3を使用しています。IMLプロシジャで、行数のみが一致する行列と列ベクトルの加算を行ないたいのですが、そのまま記述するとエラーとなります。

プログラム

```
PROC IML;
x={1 2,
3 4};
y={5,
6};

z=x+y;
```

ログに現れるエラー

```
ERROR: (execution) Matrices do not conform to the operation.
```

計算のイメージは次の通りです。

```
{1 2,3 4}+{5 5,6 6}
={6 7, 9 10}
```

どのようにしたらよいでしょうか？

**A** SAS 9.1.3またはそれ以前のリリースでは、このタイプの演算は許可されていません。たとえばREPEAT関数を使用して、列ベクトルを行列の大きさに合わせてください。

プログラム例

```
z=x+REPEAT(y,1,NCOL(x));
```

なお、SAS 9.2からはこのような演算も認められるようになりました。ただし、逆に意図しない結果が得られる可能性もあることに留意してください。

**Q** PRINTプロシジャを使って、SASデータセットの一部をランダムに表示して変数値を確認したいのですが、どのようにすれば効率的でしょうか？

**A** WHEREステートメントでRANUNI関数を使用すると、データの一部がランダムに選択されてアウトプットウィンドウに表示されます。以下のプログラムは、SASデータセットtestから全体のおよそ1%をランダムに選ぶものです。

```
PROC PRINT DATA=test;
WHERE (RANUNI(123)<.01);
RUN;
```

**Q** GLMプロシジャを用いて分析を行なっていますが、結果の表示、および出力したデータセットにおいて、効果の名前が21文字以降、表示されません。また、交互作用項では、各変数の名前がより短くなっています。長い変数の名前を用いたまま、すべて表示させるためには、どのようにしたらよいのでしょうか？

**A** 標準では、効果の名前は20文字までの表示となります。より長い変数名を用いている、また、高次元の交互作用項をモデルに含めている場合には、PROC GLMステートメントにおけるオプション、NAMELEN=オプションにて、より大きな数値を指定してください。(標準の設定はNAMELEN=20であり、最大で200まで指定することができます。)

このオプションは、GLMプロシジャの他、GENMODプロシジャ、MIXEDプロシジャ、GLMMODプロシジャ、LOGISTICプロシジャ、MIXEDプロシジャなどにも用いることができ、標準の値は同じく20となっています。なお、CATMODプロシジャにおける標準の値は24となります。

# SAS Training

## SASトレーニングのお知らせ

### 特別トレーニングコースの開催のご案内

#### 「製薬企業におけるSASプログラミング入門」コース (1日間)

日程: 2009年2月26日(木) 10:00 ~ 17:00(東京会場)  
 価格: 73,500円(税込)/ チケットのお取り扱いはありません。  
 受講対象: SASシステムによる臨床試験の統計解析またはDM業務に従事している方  
 臨床開発経験1年未満のSASプログラマーの方から、SASプログラマーを管理する方まで、幅広いレベルの方を対象としています。  
 前提知識: 下記のコースのいずれかを受講済みの方  
 ・「SASプログラミング入門」  
 ・「SASプログラミングI」  
 または、臨床開発業務に6カ月以上従事している方、さらに、下記のコースを受講済みの方がなお望ましい。  
 ・「SASプログラミングII」  
 ・「SASマクロ言語」

学習内容: 『SAS Programming in the Pharmaceutical Industry』を用いて、臨床試験におけるプログラミングのテクニックについて学習します。

コンテンツ: Chapter 1 : 環境  
 Chapter 2 : 分類変数の作成  
 Chapter 3 : データの読み込み  
 Chapter 4 : 解析用データの作成と変換  
 Chapter 5 : 表やリストの作成  
 Chapter 6 : グラフの作成  
 Chapter 7 : よくある解析の実行と結果の取得  
 Chapter 8 : データの出力  
 Chapter 9 : 臨床試験におけるSASプログラミングの将来  
 Chapter 10: リソース

テキスト: 本コースは、英語版の書籍『SAS Programming in the Pharmaceutical Industry』(SAS Press)を使用し、日本語で説明を行います。なお、補助資料等はありません。  
 【注意】受講申込者は、トレーニング申込とは別に事前にご購入ください。(上記書籍を既にお持ちの方は除きます。)購入方法は、下記『書籍購入方法』をご参照ください。

書籍購入方法: 『SAS Programming in the Pharmaceutical Industry』(SAS Press)  
 Amazon.co.jp(<http://www.amazon.co.jp/>)でご購入ください。

### 新規トレーニングコースの開催のご案内

#### SAS® Enterprise Guide®ユーザーのためSASプログラミング」コース(2日間)

日程: 2009年1月26日(月)~27日(火) 10:00 ~ 17:00(東京会場)  
 価格: 115,500円(税込)/ チケット捺印数2  
 受講対象: SAS Enterprise Guideのユーザーで、SASプログラムのコーディングや、SAS Enterprise Guideによって生成されるSASコードの修正について学習したい方  
 前提知識: 「SAS Enterprise Guideによるデータ加工とレポート作成」を受講済みか、同程度の知識のある方  
 学習内容: 本コースは、SASプログラミングへの導入となるコースで、「SASプログラミングI」コースの前提となります。  
 ・SAS Enterprise GuideとSASプログラムの連携による分析方法  
 ・SASプログラムの構造と構文  
 ・DATAステップを使用したSASデータセットの加工  
 ・DATAステップを使用した区切り記号付きテキストファイルの読み込み  
 ・DATAステップを使用したテーブルのマージ、マッチする行やマッチしない行の分離  
 ・SAS Enterprise Guideによって生成されるコードを編集した、タスク出力のカスタマイズ

コンテンツ: なぜSASプログラミングを学習するのか?

- ・コース概要
- ・SAS Enterprise Guideの基本の復習
- ・SASプログラミングを理解することの利点
- SASプログラミングの基本
- ・SASプログラム イントロダクション
- ・SASプログラミングの概念
- ・構文エラーの診断と修正
- ・SASデータライブラリへのアクセス
- ・リレーショナルデータベースへのアクセス(自己学習)
- SASデータセットの読み込み
- ・DATAステップとフィルタとクエリタスクの比較
- ・DATAステップの基本
- ・DATAステップの処理
- ・条件付き処理
- ・その他の演算子と出力形式(自己学習)
- ローデータファイルの読み込み
- ・区切り記号付きデータの読み込み
- ・データエラーの調査
- ・PDVでの値の保持(自己学習)
- SASデータセットの組み合わせ
- ・DATAステップを使用したデータのマージ
- ・データの縦結合
- ・データの追加(自己学習)
- タスクによって生成されるコードの拡張
- ・生成されたコードの理解
- ・生成されたコードとの対話
- ・ヘルプの利用(自己学習)
- 出力コンテンツの制御
- ・タイトルやフットノートの拡張
- ・コードによる実行時パラメータの利用
- ・条件によるハイライト
- 出力の種類制御
- ・アウトプットデリバリシステム
- ・HTML出力のカスタマイズ
- ・RTF出力のカスタマイズ
- ・PDF出力のカスタマイズ

### 「決定木モデリング」コース（2日間）

日程: 2009年3月2日(月)~3日(火) 10:00 ~ 17:00(東京会場)

価格: 115,500円(税込)/チケット捺印数2

受講対象: SAS® Enterprise Miner™を利用して決定木(ツリー)を構築したい予測モデリングやデータ解析に携わる分析者

前提知識: 下記2コースを受講済みか、同程度の知識のある方

- ・「SASによる統計解析I」
- ・「SASによる回帰分析I」

さらに下記のコースを受講済みか、SAS Enterprise Minerを使用した分析経験のある方

- ・「SAS Enterprise Miner 操作入門」

学習内容: 本コースは、ツリー構造の予測モデルを、また決定木の成長、剪定(枝刈り)、および評価の方法を扱っています。さらに、探索的データ解析、次元の削減、および欠損値の補完といった、ツリーの補助的な利用法が論じられます。

コンテンツ: ツリー構造のモデル

- ・分類木
- ・回帰木
- 再帰的な分割
- ・2分割と多分割
- ・分割基準
- ・欠損値
- 剪定(枝刈り)
- ・p値の調整
- ・プロフィット(利益)/ロス(損失)に関する考慮
- ・クラス確率木
- ・クロスバリデーション
- ツリーの補助的な利用法
- ・データ探索
- ・次元の削減
- ・補完
- ツリーのアンサンブル
- ・バギング(bagging)
- ・アーキング(arcing)
- ・勾配ブースティング(gradient boosting)

### 「SASプログラミング実践テクニック 2 ~処理の効率化~」 コース（1日間）

日程: 2009年3月16日(月) 10:00 ~ 17:00(東京会場)

価格: 57,750円(税込)/チケット捺印数1

受講対象: 既存プログラムの処理効率UPのテクニックを学習したいと考えている方

前提知識: 下記2コースを受講済みか、同程度の知識のある方

- ・「SASプログラミングI」
- ・「SASプログラミングII」

さらに、「SASプログラミングIII ~ 上級テクニックと効率化 ~」コースを受講いただいておりますと、本コース内容の理解に役立ちます。

学習内容: 業務上の場面を想定し、実践的なテクニックを交え学習します。  
・既存プログラムの処理速度をあげる実践的なテクニックを学習します。

コンテンツ: はじめに

- ・コース概要
- データの並べ替えテクニック
- ・並べ替えのオプション
- ・並べ替えのテクニック
- データの更新テクニック
- ・変数属性の変更
- ・データのコピー
- ・データの更新
- ・データの追加
- ・データの抽出
- ・データの結合
- データの要約テクニック
- ・MEANSプロシジャ
- ・FREQプロシジャ
- オプションテクニック
- ・メモリに関するオプション

内容は予告なく変更する場合があります。

SAS Institute Japan株式会社では、今後も多岐にわたったトレーニングコースを追加していく予定です。コース内容・日程等の詳細は、順次弊社Webサイトに公開しますので、以下のURLをご参照ください。

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

その他、トレーニングに関する情報については、上記のURL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。

#### トレーニングに関するお問い合わせ先

- TEL 03-3533-3835
- FAX 03-3533-3781
- E-mail [JPNTraining@sas.com](mailto:JPNTraining@sas.com)

# Publications

## マニュアル販売に関するお知らせ

2008年11月より弊社ホームページでお知らせしている通り、英語版 SASマニュアルおよびSAS Learning Editionリリース4.1の販売業務は、2008年12月26日をもってAmazon.co.jpへ移管しました。ご用命の際はAmazon.co.jpにてお求めください。

なお、日本語版マニュアルは、以前と同様、弊社にてご注文を受付しております。詳細は下記のURLをご参照下さい。

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

### 日本語版マニュアルに関するお問い合わせ先

- T E L 03-3533-3835
- F A X 03-3533-3781
- E-mail JPNBooksale@sas.com

## Latest Releases

### 最新リリース情報

#### PCプラットフォーム

Windows版 SAS 9.1.3 / 9.2  
64-bit Windows(Itanium)版 SAS 9.1.3 / 9.2

#### UNIXプラットフォーム

SunOS/Solaris版 SAS 9.1.3 / 9.2  
HP-UX版 SAS 9.1.3 / 9.2  
HP-UX(Itanium)版 SAS 9.1.3 / 9.2  
AIX版 SAS 9.1.3 / 9.2  
Linux(Intel)版 SAS 9.1.3 / 9.2

#### メインフレームプラットフォーム

IBM版(OS/390、z/OS) SAS 9.1.3 / 9.2  
富士通版(F4、MSP) SAS 6.09E TS470  
日立版(VOS3) SAS 6.09E TS470

9.2はIntelligence Platform製品を除く

# Information

## SAS Technical News 送付についてのご案内

SAS Technical Newsは次の方を対象にお送りしています。

- ・ SASコンサルタントとしてご登録の方
- ・ SAS Technical Newsの購読をお申し込み/いただいている方

今後SAS Technical News購読が不要の方、配信先の変更等をご希望の方は、下記URLよりお手続きください。

配信停止

<http://www.sas.com/japan/corporate/material.html>

配信先変更手続き

[http://www.sas.com/japan/sasj\\_privacy.html#inquiry](http://www.sas.com/japan/sasj_privacy.html#inquiry)

## SAS Technical News Winter 2009

発行  
SAS Institute Japan株式会社

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



SAS Institute Japan株式会社

[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  
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