SAS Technical News

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



SUMMER 2011

特集	0
SAS® 9.3の実行環境と	
関連ドキュメントについて	

特集2 …………06

SAS 日本語版 新刊マニュアルのご紹介

SAS Academic News ····· 09

- 事例紹介
- コラム「SAS四方山話」

Q&A · · · · ·	• • • • • • • • • • • • • • • • • • • •	16
SASトレーニ	ングのお知らせ	19
旦並!!!! フ	AR ±0	

Requirements for SAS 9.3

SAS® 9.3の実行環境と 関連ドキュメントについて

2011年7月にSASの最新バージョンである、SAS 9.3 が米国でリリースされました。これからSAS Technical Newsでは、何回かに分けてこの新しいバージョンの 変更点や新機能などをご紹介していきます。一回目の 今回は、導入を検討するにあたって、事前に確認いた だきたいSAS 9.3の実行環境についてご紹介します。 なお、ここに記載している情報は2011年7月初めの 時点のものであり、今後、新しいプロダクトの提供などに 合わせて、変更・追加される可能性があります。実際に 導入をご検討の際は、SAS インストールセンターから最新 のシステム必要条件のドキュメントなどをご確認ください。



1. SAS 9.2からの変更点

次の2つのOSがサポート対象から除外されましたので、 ご注意ください。

- > OpenVMS Itanium
- > Microsoft Windows Itanium

加えて、Webアプリケーションサーバーなどを配置するミド ル層においては、32-bit 版 Windows がサポート対象外 になりましたので、ご注意ください。これにより、ミドル層は、 64-bit 環境のみのサポートになります。

一方、SAS Metadata Serverのサポート対象に64-bit 版z/OSが追加されました。

2. SAS 9.3 Foundation O 対応オペレーティングシステム

続いて、SAS 9.3 Foundationの対応OSについて、OSごとに 記載します。

Operating System

SAS 9.3の 対応オペレーティングシステムと変更点

まず、SAS 9.3の対応オペレーティングシステム(以下OS)について ご紹介します。

具体的な対応OSの説明を始める前に、SAS 9.2からの変更点について 簡単に記載します。

UNIX 版

> HP-UX (PA-RISC): HP-UX 11i Version 3(11.31), Update 5以上。 PHNE 41004 11.31 cumulative ARPA Transport patch 以降のパッチ

DVDメディアからSASをインストールする場合は、さらにPHCO_41366 パッチをOSに適用する必要があります。

また、HP-UX 11.31を使用する場合は、Atomic APIsプロダクト B.11.31.0803.01が必要です。

> **HP-UX** (**Itanium**) : HP-UX 11i Version 3 (11.31)、Update 5以上。 PHNE_41004 11.31 cumulative ARPA Transport patch 以降のパッチを適用。

DVDメディアからSASをインストールする場合は、さらにPHCO_41366 パッチをOSに適用する必要があります。

また、HP-UX 11.31を使用する場合は、Atomic APIsプロダクト B.11.31.0803.01が必要です。

- > AIX (Power CPU、または、RS64アーキテクチャ): AIX Release 6.1以上。
- > **Solaris (SPARC)**: Solaris 10 Update 8 for UltraSPARC III 以上をサポートするすべてのモデル。SAS 9.3 は、SPARCV9 CPUアーキテクチャが必須です。
- > Solaris (x64): Solaris 10 1/06以上。

なお、Solaris 10 06/06を使用する場合は、ディスプレイマネージャモードでのエラーを回避するためパッチ119281-06が必要です。

また、LDAP認証を使用する場合は、120037-09以降のパッチを適用してください。

SASでは、Oracle 社が推奨している最新のSolaris Patch Cluster for x64 Solaris 10を適用することを推奨しています。

> **Linux (Intel**アーキテクチャ): RedHat Enterprise Linux 5 Update 4、Red Hat Enterprise Linux 6、Novell SuSE Linux Enterprise Server 10 SP3 および Novell SuSE Linux Enterprise Server 11。

RedHat Enterprise Linux 5以降を使用する場合は、ライブラリlibxp.soのインストールが必要です。RedHat Enterprise Linux 6を使用する場合は、libXext-1.1-2.el6.i686パッケージのインストールが必要です。

さらに、SAS 9.3 Foundationでは、Linux 5、6ともにGLIBC 2.4が必須です。

> Linux (x64): RedHat Enterprise Linux 5 Update 4、Red Hat Enterprise Linux 6、Novell SuSE Linux Enterprise Server 10 SP3 およびNovell SuSE Linux Enterprise Server 11。

RedHat Enterprise Linux 5以降を使用する場合は、ライブラリlibxp.soのインストールが必要です。RedHat Enterprise Linux 6を使用する場合は、libXext-1.1-2.el6.i686パッケージのインストールが必要です。

さらに、SAS 9.3 Foundationでは、Linux 5、6ともにGLIBC 2.4が必須です。

Windows版

> Windows:32-bitおよびx64版 Microsoft Windows Server 2008 (Standard、Enterprise、Datacenter Edition)、x64版 Microsoft Windows Server 2008 R2 (Standard、Enterprise、Datacenter Edition)、32-bit版 Microsoft Windows Server 2003 SP1 (Standard、Enterprise、Datacenter Edition)、Microsoft Windows Server 2003 for x64 (Standard、Enterprise、Datacenter Edition)。

z/OS版

> **z/OS**: z/OS, Version 1, Release 10以上。



- クライアント層でサポートしているOSには、次のようなものがあります。
- > Windows XP Professional SP2
- > Windows XP Professional for x64 (32-bit 互換モードも含む)
- > Windows Vista
- > Windows Vista for x64 (32-bit 互換モードも含む)
- > Windows 7
- > Windows 7 for x64 (32-bit 互換モードも含む)

原則、Home Editionはサポート対象外です。ただし、Academic Packageでは例外もありますので、その場合は、SASテクニカルサポートまでお問い合わせください。



Middle Tier

SAS 9.3 ミドル層: Webアプリケーションサーバーについて

SAS 9.3でサポートしているWebアプリケーションサーバーの対応バージョンは次の通りです。

- > IBM WebSphere 7以上
- > BEA WebLogic 11gr1 (10.3.x) 以上
- > JBoss Web Application Server 4.2.3 GAまたはEAP 4.3以上なお、前述のとおり、ミドル層では64-bit版OSであることが必須になりましたので、ご注意ください。

SAS 9.3でサポートするWebアプリケーションサーバー

os	バッァ 数	JREのビッァ数	JBoss EAP 4.3 (4.2.3 Community Edition) の使用	Oracle WebLogic (WLS) 10.3.x の使用	IBM WebSphere (WAS) 7 の使用
HP-UX Itanium	64-bit	64-bit	0	0	×
x64 Linux	64-bit	64-bit	0	0	0
Solaris (SPARC)	64-bit	64-bit	0	0	0
Solaris 10 x64	64-bit	64-bit	0	0	0
AIX(POWER)	64-bit	64-bit	0	0	0
Windows x64 (x86-64)	64-bit	64-bit	0	0	0
z/OS	31-bit	31-bit	X	X	0





SAS 9.3のJavaの必要条件

SAS 9.3からの Java の必要条件の主な変更点は、下記のとおりです。

- > Oracle 社から提供される JRE を使用します。
- > これまでは、Java 1.5.xをサポートしていましたが、Java 6のサポー トに変わります。

下記の表にOSごとの必要条件を記載します。

SAS 9.3でサポートするJRE

OS	サポートするJREのバージョン
HP-UX PA-RISC	6.0.07
HP-UX Itanium	6.0.07
Linux(32-bit、x86版)	1.6.0_21
Linux(64-bit、x86-64版)	1.6.0_21
AIX	Java 6 SR8 FP1 + iFix IZ77436
Solaris(64-bit、SPARC)	1.6.0_21
Solaris(64-bit、x86-64版)	1.6.0_21
Windows(32-bit、64-bit版両方)	1.6.0_24 (32-bit版JREが必須。 64-bit版OS使用の場合も 32-bit版JREを使用してください。)
z/OS	Java 6 31-bit SDK

SAS 9.3でサポートするJDK

os	サポートするJDKのバージョン
HP-UX Itanium	6.0.07
Solaris(64-bit、SPARC)	1.6.0_21 JRockit 6 Update 24 R28.1.3 (WebLogicを使用の場合のみサポート)
Solaris(64-bit、x86-64版)	1.6.0_21
AIX	Java 6 SR8 FP1 64-bit(6.0.0.215)
Linux(64-bit、x86-64版)	1.6.0_21 JRockit 6 Update 24 R28.1.3 (WebLogicを使用の場合のみサポート)
Windows(64-bit、x86-64版)	1.6.0_24 JRockit 6 Update 24 R28.1.3 (WebLogicを使用の場合のみサポート)
z/OS	Java 6 31-bit SDK



SAS 9.3で対応するWebブラウザは、下記のとおりです。

- > Internet Explorer 7
- > Internet Explorer 8
- > Firefox 3.6

サポートするWebブラウザとOSの対応については、下記の表をご参照 ください。

Webブラウザの 種類	os	Webブラウザの バージョン
	Windows XP SP2/SP3 (32-bitまたは、64-bit x64)	7,8
	Windows Server 2003 (32-bit, 64-bit x64)	7.8
Internet Explorer	Windows Vista (32-bit または、64-bit x64)	7,8
	Windows Server 2008 (32-bit または、64-bit x64)	7,8
	Windows 7 (32-bit または、64-bit x64)	8
	Windows Server 2008 R2	8
Mozilla Firefox	Linux 32-bit (x86)	3.6
	Windows XP SP2 (32-bit または、64-bit x64)	3.6
	Windows Vista (32-bit または、64-bit x64)	3.6



SAS® Add-in for Microsoft Office (AMO) など一部 BI 関連のクライ アントを使用するお客様は、サポートするMicrosoft Office のバージョ ンの確認も必要です。SAS 9.3でサポートするMicrosoft Officeのバー ジョンは、下記のとおりです。

- > Microsoft Office 2003
- > Microsoft Office 2007
- > Microsoft Office 2010





SAS 9.3でサポートするDBMS

SAS 9.3で対応する主な DBMS のバージョンは、下記のとおりです。 SAS/ACCESSプロダクトなどを活用してDBMSからのデータを活用 している場合は、バージョンをご確認ください。ここに記載されていな いDBMSに接続するSAS/ACCESSプロダクトをご利用の場合は、 「SAS 9.3 Foundationシステム必要条件」のドキュメントに記載されて いる情報を参照してください。

- > DB2: DB2 Universal Database, Version 8.1 FixPak 18 以上 (32-bit または、64-bit ライブラリ。OS に依存。)
- > MySQL: MySQL Client 5.1または、MySQL Client 5.5。MySQL Server 5.0以上。
- > Netezza: Netezza Version 4.0以上。Version 5.0以上の使用を推奨。 使用するNetezzaのバージョンに対応するNetezza ODBCドライバも 必要。
- > Oracle: Oracle Release 10g以上。
- > **Sybase**: Sybase Open Client SDK Software, Release 12.52以上。
- > Teradata: Teradata Database 12以上(Temporal機能を使う場合 は Teradata Database 13.10以上が必要)。 Teradata CLIv2 client libraries, TTU12以上(Temporal機能を使う場合はTTU 13.10以上

加えて、SASでは次のソフトウェアの使用を推奨しています: Teradata FastExport utility, TTU12以上、Teradata MultiLoad utility, TTU12 以上、Teradata Parallel Transporter API, TTU12以上。

SAS 9.3対応 インストールセンターについて

SAS 9.3 のシステム必要条件をまとめた『システム必要条件』、イン ストール手順や構成手順を記載した『SAS Deployment Wizard ユーザーガイド』、『クイックスタートガイド』、『設定ガイド』などの ドキュメントは、これまでと同様に下記のURLのSASインストール センターに掲載します。

SAS インストールセンター

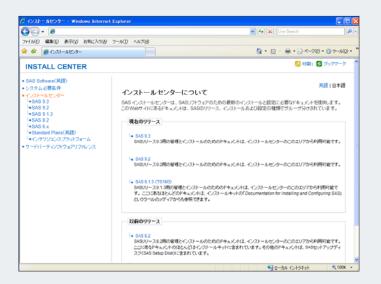
http://www.sas.com/japan/service/documentation /installcenter/index.html



All the documents

are available at **

the SAS Install Center.



なお、スクリーンショットは現在準備中のものです。実際の開設時には レイアウトなどが変更される可能性もありますので、あらかじめご了承 ください。

今回、この特集記事でご紹介した、SAS 9.3の実行環境を確認する場合に参照する『システム必要条件』、インストール手順や構成手順を記載した、『SAS Deployment Wizardユーザーガイド』、『クイックスタートガイド』、『設定ガイド』などのドキュメントは、下記の「SAS 9.3 (TS1M0)ドキュメントページ」に掲載されます。

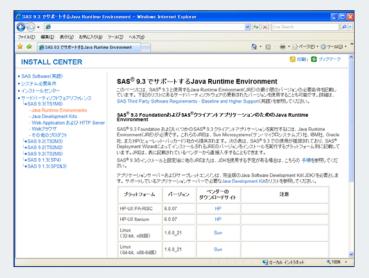


このページを参照する場合は、まず、インストールセンターのトップページ中央部分の「現在のリリース」からSAS 9.3を選択します。次に表示されるページの「ドキュメント一覧」から[SAS 9.3(TS1M0)のドキュメント]を選択してください。

また、Webアプリケーションサーバーや Java の必要条件などは、次の「サードパーティソフトウェアリファレンス」のページに掲載予定です。 このページを参照するには、インストールセンターの左側のメニューから「サードパーティソフトウェアリファレンス」を選択し、さらに「SAS 9.3 (TS1M0)」を選択します。



その後、参照したい情報のページ(たとえば、JREなど)を選択します。



おわりに

今回は、SAS 9.3の実行環境とその最新情報を掲載するSAS インストールセンターの利用方法について、ご紹介しました。新規にSAS 9.3の導入を検討する場合、また、現在ご利用のSASを最新版にアップグレードする場合にどのような環境が必要なのか考える際の目安となれば幸いです。

今回のリリースは、SAS 9.3の初期リリースです。これから、秋、冬と新しいプロダクトやソリューションがリリースされる予定です。

次回以降の SAS 9.3 の特集では、新機能などのご紹介をする予定です。どうぞお楽しみに。



特集2

SAS日本語版 新刊マニュアル 『SAS®認定プロフェッショナルのための **Base Programming for SAS®9** 完全ガイド』のご紹介

2011年7月1日より待望の日本語版マニュアル、

[特集2] SAS 日本語版 新刊マニュアルのご紹介

『SAS® 認定プロフェッショナルのための Base Programming for SAS® 9 完全ガイド』を発売開始しました。

今回は、こちらの書籍の魅力をご紹介するとともに、「SAS Technical News」読者のみなさま限定の特別価格でご提供いたします。



SAS® 認定プロフェッショナルのための Base Programming for SAS® 9 完全ガイド



読者限定キャンペーン -

通常価格: 18,900円(税込)



特別価格: 17,010円(税込)

- 申込期限: 2011年9月30日(金)受付分まで
- 申込方法:下記URLから本キャンペーン専用の申込書 をダウンロードの上、ご利用ください。
- ※ 専用申込書以外の用紙でお申込みの場合は、割引が 適用されませんのでご注意ください。

SAS 新刊マニュアル申込書

www.sas.com/jp/manual/PrepGuide_Tech News.pdf

- ISBN: 978-4-87235-163-7
- 出版元: SAS Institute Japan株式会社
- サイズ: A4版
- ページ数:784ページ
- 書籍詳細はこちら:

www.sas.com/jp/manual/bp9_guide.html (目次・本文なども一部閲覧できます)

本書は、SAS初心者が「SAS Base Programming for SAS® 9」 試験*¹⁾ (以下、Base 試験) 合格に向け、必要なスキルを身につける ための最適な自習書です。

Base試験は、SASの基本概念からデータ読み込み、加工、分析、レポーティングといった一連のSASを使用した処理の理解度を問う試験です。 試験範囲は多岐にわたるため、本書は22章、784ページにもおよび、充実した内容をカバーしています(図1:章立てと主な学習内容)。 また、本書は認定試験対策のみならず、SAS初心者の入門書として、SAS経験者の日本語マニュアルとしても活用できます。以降ではそれ

Base認定試験の対策本として

ぞれの用途別の活用術をご紹介します。

Base 試験の合格には、幅広い網羅的な知識・経験が求められます。 対象となる SASトレーニングは、合計 6 日間分にわたります。本書はこれらの試験範囲の要点が詰まっている、お得な書籍と言えるでしょう。

・直前対策として

各章の後半には基本的な知識を問うクイズが用意されています。直前 対策として使用する場合は、まず初めにクイズを解き、解説を参照して 自分の弱点を洗い出すことができるでしょう。必要に応じて詳細情報を 確認して効率的に試験準備が行えます。

より応用的な練習問題にチャレンジする場合は、別売のe-learning 教材*²⁾ との併用をおすすめします。

・初心者から認定資格者へ

本書を使用すると、SASやプログラミングの経験がまったくない方でも、 着実に知識を積み重ねることができます。SAS ウィンドウ環境の基本 操作や環境設定についても詳細に解説しているため、本書中のサン プルプログラムを入力して動作確認を行うことで経験を積むことも可能 です。また、基本概念→データ読込・加工とレポーティングの基本→ レポーティングの応用→データ加工の応用というバランスの良い構成 で、着実に知識が広がり、全体感をとらえることができます。

プログラミング未経験からBase 試験に合格された方のお話を伺うと、 平日に2時間程度の学習を3カ月継続するくらいが学習時間の目安の ようです。たとえば「初めの2カ月で毎週3章分の学習を行う」など、章の ペース配分を決めながら計画的に学習を進めることも可能になります。

初心者向けの自習書として

認定試験を受験しない方でも、自習書として役立ちます。DATA ステップの理解において、特に重要な内部動作についてここまで詳しく解説している日本語の書籍は他に見当たらないと思います。

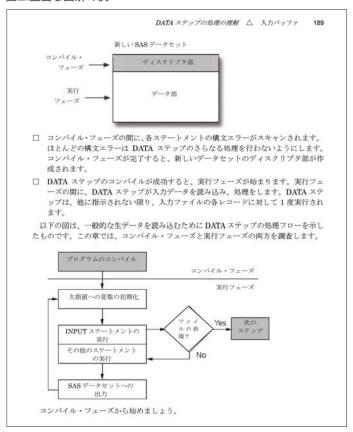
図1:章立でと主な学習内容

章	基本概念	データ読み込み	データ加工	レポート作成
1.基本コンセプト	基本用語·基本操作			
2.ファイルの参照と設定のオプション	ライブラリ			
3.SAS プログラムの編集とデバッグ	ログ、デバッグ			
4.リストレポートの作成	ラベル、出力形式		SORT	PRINT
5.生データファイルとExcelワークシートの読み込み	SAS日付定数	固定長、Excel	割り当て、サブセットIF	
6.DATA ステップの処理の理解	コンパイル、実行	生データ(固定長)	条件分岐(IF-THEN)	
7.ユーザー定義の出力形式の作成と適用	独自の出力形式	生データ(固定長)		FORMAT
8.拡張リストレポートと要約レポートの作成				REPORT
9.記述統計量の作成				MEANS, FREQ
10.HTML 出力の作成	ODS			HTML
11.変数の作成と管理		生データ(固定長)	累積合計、条件分岐	
12.SAS データセットの読み込み		SAS データセット		
13.SAS データセットの組み合わせ		SAS データセット	マージ、APPEND	
14.SAS 関数でのデータの変換		SAS データセット	関数	
15.DO ループでのデータの生成		SAS データセット	ループ	
16.配列での変数の処理		SAS データセット	配列	
17.固定長フィールドの生データの読み込み		生データ(固定長)		
18.フリー・フォーマット・データの読み込み		生データ(区切文字)		
19.日時値の読み込み	SAS日付			
20.複数のレコードからの単一オブザベーションの作成		生データ(複数読込)		
21.単一のレコードからの複数のオブザベーションの作成		生データ(複数出力)		
22.階層型ファイルの読み込み		生データ		



6章では「DATA ステップの動作がコンパイルフェーズと実行フェーズに 分かれる」「コンパイルフェーズで SAS データセットファイルのディスクリ プタ部が作成される」「実行フェーズでデータ部が作成される」などの 重要な概念と実行時の動作を、図を交えながらわかりやすく解説して います(図2)。

図2:豊富な図解の例



初心者の方はまず6章までの基本部分を理解しましょう。その後、 必要に応じて後半の章を学習すればよいでしょう。例えば、生データ ファイルの読み込みを行うのであれば、後半の章のみをピックアップ して学習できます。

経験者向けの日本語マニュアルとして

本書は、SAS経験者の方の日本語マニュアルとしても有効です。図3 に示すようなステートメント、関数、プロシジャの使用方法を調べること ができます。

また基本的な使用方法だけでなく、少し高度なデータ読込なども解説 しています。例えば、SET ステートメントの POINT= オプションを使用し てオブザベーション番号を指定したSASデータセット読み込みを行う 方法などは、5章で解説しています。

*1) SAS® 認定プロフェッショナル SAS Base Programmer for SAS® 9 この資格は、プログラマー、分析担当者、データ管理者や、クエリや分 析を行うためにデータアクセスや管理用のSASプログラムを記述して いる方に最適です。この資格は、その他多くの役割を担うための基盤 となる知識として適切です。

詳細はこちら: www.sas.com/jp/training/certify/bp9.html

*2) e-Learning 教材

「SAS 認定プロフェッショナル模擬問題: SAS Base Programmer for SAS® 9 」「SAS Base Programming for SAS 9」の模擬テストとして 利用でき、試験と同様の知識と技能をテストします。

詳細はこちら:

www.sas.com/jp/training/e-learning/prep_base.html

以上、本書の活用方法をご案内しました。

皆様のお手元に置き、試験勉強に、自習書に、マニュアルに、幅広く ご活用ください。

日本語マニュアルに関するお問い合わせ先

エデュケーション部 マニュアル販売係

T E L: 03-6434-3690 F A X: 03-6434-3691

E-mail: JPNBooksale@sas.com

図3:本書で解説している主な言語構成要素

分類	要素
ステートメント	DATA、SET、WHERE、MERGE、INFILE、INPUT、DATALINES、RUN、KEEP、DROP、FORMAT、LABEL、LENGTH、RETAIN、BY、IF-THEN-ELSE、SELECT、WHEN、OTHERWISE、DOグループ、DOループ、DO UNTIL、DO WHILE、END、ARRAY、TITLE、FOOTNOTE、ODS、OPTIONS、FILENAME、LIBNAME、OUTPUT、DELETE、FILE、PUT、STOP
変数	_ERROR_,_N_, First., Last., IN=, END=
関数	DATDIF, DATE, DAY, INTCK, INTNX, QTR, MDY, MONTH, TIME, TODAY, YEAR, YRDIF, CATX, FIND, INDEX, LOWCASE, SUBSTR, SCAN, TRANWRD, TRIM, UPCASE, PROPCASE, DIM, PUT, INPUT, INT, MAX, MEAN, MIN, STD, SUM, ROUND, VAR
プロシジャ	PRINT、FREQ、MEANS、SUMMARY、REPORT、CONTENTS、SORT、FORMAT、APPEND

09



SAS アカデミック・ニュース

cademic News

事例紹介、第5章 SAS® Enterprise Guide 記述統計メニューで実行する各種統計の 分布と分割表分析についてご紹介いたします。

前号では記述統計のメニューからデータの傾向を見るという解説でした。

今号では、さらに一歩踏み込み分布図や分割表分析からデータの分布状態や 特定の統計分布に合致しているか等について解説しております。

SAS四方山話では、第2話研究開始に向けての目的や研究方針の決定等、 論文を作成する上での論文全体の設計についてご紹介いたします。



事例紹介

コラム 「SAS四方山話」

事例紹介

高柳 良太 國學院大學 経済学部および人間総合科学大学 人間科学部 兼任講師

第5章 SAS® Enterprise Guideでの分布と分割表分析

今回は、SAS Enterprise Guide (以下、 EG)の「記述統計」メニュー内で実行でき る各種統計量の算出機能のうち、「分布」と 「分割表分析」の説明をしたいと思います。

> 記述統計(B) - グラフ(G) - 分析(Z) -| リスト(L)... 要約統計量ウィザード(M)... Z Σ 要約統計量(S)... 関 集計表ウィザード(B)... 集計表(T)... データの特性分析(出)... 分布(D) 一元度数表(O)... **分割表分析(A)...**

記述統計メニュー

前回は、「記述統計」のメニューを使って データの傾向を把握するということを説明 しました。基本的には、今回もその続きの 話となります。ただし、今回は傾向を把握 するということに加えて、そのデータが何 かと同じなのか、またはカテゴリ別にみた ときに違いがあるのかという「検定」という ものが一部入ります。

学生や統計に慣れていない人に、統計 解析について説明していると、この「検定」 の話が一つの山となります。検定は、正確 には統計的仮説検定や統計的検定と呼 びますが、統計や確率の独特な考え方を 象徴しているものの一つと言っていいで しょう。どれくらい難しいかというと、書店の 統計書籍のコーナーには「わかりやすい」 をキャッチフレーズにした「検定」の書籍 がたくさんあります。たくさんあるということ は、それだけわからない人がたくさんいて、 わかりやすく説明した本が少ないというこ とだと思います。

私も統計の本を書いているので、人様の ことを言える立場ではないのですが、統計 解析についてわかりやすく説明するという のは、なかなか難しいことです。実際の データを使って話ができればいいのです が、書籍で説明する場合は説明用のデー タを用意して、それで説明をするので、なか なか実感を持って理解していただけないと いう問題があります。

今回はニュースレターの1コーナーなの で、細かく説明する余裕がないので非常に 申し訳ないのですが、とりあえず「そういう ものだ」という風に思ってください。こういう ことを書くと、統計を専門にしている先生 からお叱りを受けるのですが、とりあえず 「そういうものか」ということで進まないと、 止まったままになってしまいます。データ分 析を必要としている方は、大抵は統計の 専門ではなく、マーケーティングや金融な ど、統計以外の分野の専門の方だと思い ます。もちろん知識はあって邪魔になるも のではありませんが、目的がデータの解 析である以上、多少の細かいところは目を つぶっても、走り抜けることが大事だと思う のです。

今回は検定の話が出てきますが、「無理 そうだ」と思ったら「そういうものか」と思っ て走り抜けてください。ただ、EGで「そうい うことができる」ということは知っておいてく ださい。そうすると、きっといつかどこかで 何かの役に立つと思います。

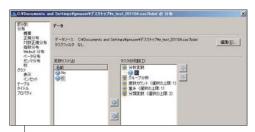
分布





「分布」は何をするのかというと、文字通 りデータがどのような分布をしているか調 べるところです。データの分布の状態を知 ることの他に、特定の統計分布に合致して いるかどうかを調べることができます。統 計解析では特定の統計分布に合致してい る(特定の分布と見なすことができる)こと を前提とした手法があり、そのようなものを 利用したい場合にはその分布と見なせる かどうかを調べます。

まずは、データがどのような分布をしている かを調べましょう。ここでは、新入社員に実施 した英語のテストの得点データを使用します。





上 分布のデータ / 下 分布のグラフ

他のEGの分析と同様に、「データ」で 分析したい変数を「分析変数」に指定しま す。データの分布傾向を把握する場合は、 「グラフ」の「表示」でグラフ出力の種類を 選択するとよいでしょう。「グラフ」は以下 の5種類が選択できるようになっています。

- ●ヒストグラム
- ●確率プロット
- **●**00プロット
- ●箱ひげ図
- ●テキストベースのプロット

テキストベースのプロットは、オブザベー ション数(データ数)に応じて幹葉図また は棒グラフ、および箱ひげ図、正規確率プ ロットをテキストベースで出力します。箱ひ げ図と確率プロットは出力メニューが他に もありますが、幹葉図はここでしか出力で きません。

データの傾向を把握する場合は、ヒスト グラム、箱ひげ図、テキストベースのプロッ トをチェックすればよいでしょう。箱ひげ図 はテキストベースでも出力されるので、場 合によっては指定しなくてもよいでしょう。

分布を実行すると、基本統計量が出力さ れます。

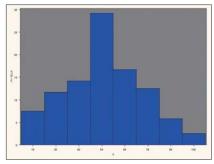
「分布」の出力



基本統計量は、平均、中央値、最頻値、 標準偏差、分散、範囲、四分位範囲が出 力されます。四分位範囲は、75パーセンタ イルの値から25パーセンタイルの値を引 いたものです。50パーセンタイル(中央値) を中心に前後25%、全体の50%の値の 範囲ということになります。

基本統計量出力の下には、平均、標準偏 差、分散の95%信頼区間が表示されます。 元の分布が決まっていて、今回のデータが それに当てはまるか調べる際に使用します。

その下の「位置の検定」は、平均値が0で あるかどうかの検定を行った結果です。 特に必要がないのであれば、メニューの 「テーブル」で「位置の検定」のチェックを 外すと出力されません。初期設定では出力 されるようになっています。



ヒストグラム

ヒストグラムは、データを等間隔の区間 に分け、その区間内にどれだけのデータが あるかをグラフ表示したものです。データ の大まかな分布状況を把握するのに役立 ちます。

「テキストベースのプロット」を選択する と、以下のような出力が得られます。

左から幹葉図、箱ひげ図、正規確率プ ロットです。

幹葉図はこの場合「幹」が10の位、「葉」 が1の位を表しています。例えば、一番下2 行で説明しましょう。

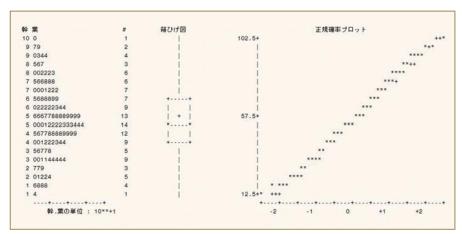
1	6888	4
1	4	1

下の行で説明すると、一番左の1は幹で 10の位で、次の4は1の位です。ですからこ こには14というデータがあることになりま す。一番右はこの区間のオブザベーション 数ですから、この区間には14が一つあると いうことになります。同じように考えると、そ の上の区間は16,18,18,18という4つのデー タがあることになります。

幹葉図は、ヒストグラムを横にしたような 状態なので、データの散らばり具合と各区 間のデータの状態を一度に把握することが できます。今回は新入社員120名のデータ ですが、もっと多いデータや、特定の区間に データが集中している場合などは幹葉図で は見づらいので棒グラフが出力されます。

箱ひげ図は、箱とひげ(棒線)からできて います。目盛りは幹葉図の目盛りを使いま す。ひげの上下の端が最大値と最小値に なり、箱の上下が75パーセンタイルの値と 25パーセンタイルの値になり、箱が四分位 範囲ということになります。真ん中の線が 中央値で50パーセンタイルになります。 箱の中の+は平均値の位置です。

正規確率プロットは、データが正規分布 に従っていると仮定した場合に、直線状に なるグラフです。



テキストベースのプロット

分布の適合度検定 正規分布を例に



「分布」の「概要」

集めたデータが特定の分布に従って いるかどうかを知りたい場合、「分布」の 「概要」で分布を指定することで、当ては まっているかどうかを知ることができます。 利用可能な分布は、以下のとおりです。

- ●正規分布
- ●対数正規分布
- ●指数分布
- ●Weibull分布
- ●ベータ分布
- ●ガンマ分布
- ●核

「核」はノンパラメトリックな各密度推定 で、初期設定は漸近平均積分事情誤差を 最小にする区間推定となっています。

ここでは、正規分布を例に簡単な例を説 明したいと思います。「分布」の「概要」の 「利用可能な分布」で「正規分布」を指定 します。指定はこれだけです。実行をする と、前述の分布に加えて、以下のような 出力がされます。

正規分	布のパラ	ラメータ
バラメータ	記号	推定値
平均	Mu	54.9
標準偏差	Sigma	20.21789

正規	分布へ	の連合度検知	Ē	
検定	統計量 p値			
Kolmogorov-Smirnov	D	0.04464931	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.02885975	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq	0.23466563	Pr > A-Sq	>0.250

正規分布の分位点			
	分	位点	
バーセント	観測値	推定値	
1.0	16.0000	7.86616	
5.0	20.5000	21.64453	
10.0	28.0000	28.98973	
25.0	41.5000	41.26324	
50.0	54.0000	54.90000	
75.0	69.0000	68.53676	
90.0	82.0000	80.81027	
95.0	91.5000	88.15547	
99.0	99.0000	101.93384	

正規分布への適合度検定の出力

最初の「正規分布のパラメータ」は、平均 値と標準偏差が出力されています。こちら では推定値となっているのは、母集団の値 を手持ちのデータ(標本)から推定してい るという考えだからです。母集団の推定値 という扱いですが、標本の実際の値です。

次の「正規分布への適合度検定」は、 データが正規分布か(正確に言うと、デー タの母集団が正規分布に近似していると見 なせるかどうか) を検定しているところです。

紙面が少ないのでものすごく省略します が、統計的検定は「比べたいものと差がな い」という仮説について検証するものです。 この場合はデータが正規分布と差がない (違わない)という仮説(正確には帰無仮 説)に基づいて計算を行っています。この 「差がない」「帰無仮説」については、学生 に教えていても「わからない!」が連発さ れるところなので気になるところではありま すが、とりあえずは「違いがないことを調 べようとしている」と思っておいてください。

正規分布への適合度検定、通称では 正規性の検定といわれますが、正規性の 検定はいくつか種類があります。メジャー どころは一番上の「Kolmogorov-Smirnov」 (コルモゴロフ-スミルノフ)の検定です。 右側のp値というところが「>0.15」となって いますが、これは「コルモゴロフ-スミルノフ 検定の計算を行って統計値を得たが、この データが正規分布と等しいと仮定した場 合に、その値は0.15(15%)以上の確率で 得られるものである」ということです。

統計的検定を行う場合、この確率が 0.05(5%)を下回っていた場合に、「同じ であるということで話を進めてきたのは 間違いだった」という判断を下すことにな ることが多いのですが、今回はそれより大 きな値となりそうなので、今回のデータは 正規分布と見なしてもいいのではないかと 判断することになります。

その下は「正規分布の分位点」で、各パー センタイルの実測値と、それが正規分布 だった場合の推定値です。

なお、正規性の検定は他からも指定する ことができます。メニューの「テーブル」で 「正規性の検定」にチェックを入れます。



「テーブル」で「正規性の検定」をチェック

ここでチェックを入れると、「分布」の「概 要」で指定をしなくても以下の出力を得る ことができます。

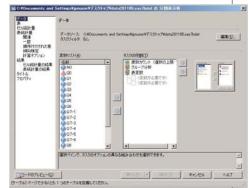
正規性の検定				
検定	15	計量	p fi	ı
Shapiro-Wilk	W	0.987406	Pr < W	0.3334
Kolmogorov-Smirnov	D	0.044649	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.02886	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.234666	Pr > A-Sq	>0.2500

「テーブル」で「正規性の検定」をチェックした場合の出力

「テーブル」で「正規性の検定」をチェッ クした場合は、「Kolmogorov-Smirnov」 だけでなく「Shapiro-Wilk」(シャピローウィ ルク)検定も出力されます。シャピローウィ ルク検定の方が、コルモゴロフースミルノ フ検定よりも新しい検定で、よく利用されて います。

分割表分析





上「記述統計」メニューの「分割表分析」

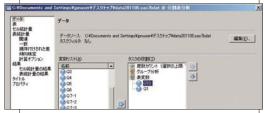
下 分割表分析のメニュー

データをある属性ごとに分割して集計す ると、より特性などが把握しやすいことが あります。この、データをある属性ごとに分 割した表のことを分割表とかクロス集計表 とよびます。難しく考える必要はなく、男女 別や購入者・未購入者別にデータを集計 したいというものです。

分割表はExcelでもピボットテーブルを 使え出力ができますが、集計を行うにマニ アックな知識が必要となります。EGの場 合は、集計したい変数を指定するだけで 簡単に分割表を作成することができます。 今回は、前回の説明と同様に学生調査の データで説明しましょう。

学生調査に、性別(1. 男性 2. 女性)と 虫歯の有無(1.虫歯になったことがない 2. 治療済み 3. 治療していないものがある) というデータがあります。男女別に、虫歯の 有無を集計してみたいと思います。

分割表分析も、他のEGの分析と同様に、 「データ」で分析したい変数を指定します。 分割表分析の場合は、「表変数」に指定し ます。



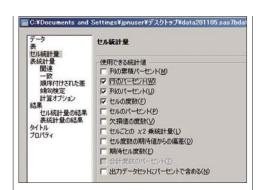
「分割表分析」の「データ」分析したい変数を「表変数」に指定

今回はQ1が「性別」で「Q10」が「虫歯 の有無 | なので、Q1とQ10を表変数に指 定します。次に「表」で行側(表側)にQ1 (性別)、列側(表頭)にQ10(虫歯)を指 定します。



「表」で分析する変数を配置

次に、「セル統計量」で「行のパーセント」 「列のパーセント」「セルの度数」にチェッ クをつけます。



「セル統計量」で出力を指定

実行をクリックすると、以下の出力を得る ことができます。



分割表分析の出力

X二乗検定

分割表分析を使うと、今まで説明してき たように、属性のカテゴリ毎の度数を求 めることができます。この例でいうと、Q1 の性別ごとに「行のパーセント」を見ると、 男子学生はなんと44.6%が3の治療して いない歯があると答えています。女子学 生では治療していないのは2名のみで 8.3%です。大学生はいちいち親が面倒を みない年齢なのですが、新しく虫歯になっ たのか、過去の詰め物が取れてしまった のか、いずれにしろ見落とせない数字で す。ところで、このデータから男子学生の 方が放置している虫歯が多いと言えるの

そういうときにも「検定」を使います。こ のような場合は、 $[\chi$ 二乗検定」を利用し て、男子学生と女子学生で虫歯の有無の 比率に違いがあるのかどうかを調べま す。比率自体は分割表で出ているので、 後は検定をするだけです。

検定は、「分割表分析」メニューの「表 統計量」の「関連」で、「関連性の検定」 から「 χ 二乗検定」をチェックします。



「表統計量」の「関連」で、

「関連性の検定」から「x二乗検定」をチェック

実行すると、分割表の下に以下のような 検定結果が出力されます。

統計量	自由度	値	p fil
カイ 2 乗値	2	9.9223	0.0070
尤度比カイ2 乗値	2	11.5558	0.0031
Mantel-Haenszel の力イ2 乗値	1	7.5468	0.0060
ファイ係数		0.3522	
一致係数		0.3322	
Cramer の V 統計量		0.3522	

χ二乗検定の出力

検定の結果の見方は、先ほども説明し ましたとおり「比較したいものに差がない」 という仮説(帰無仮説)を検証するように できています。今回でいうなら、「男子学生 と女子学生の間で、虫歯の有無に違いは ない」というのが帰無仮説となります。

ちなみに帰無仮説は「きむかせつ」と読 みます。余談ですが、「コルモゴロフースミ ルノフ検定しのように、統計用語には外国 人の統計学者の名前がついているものが 多くあります。余談ですが、一つ笑い話が あります。大学で授業をしていたときに、 「先生、『きむかせつ』って韓国の人が考 えたのですか?」という全くもって予期せぬ 質問を受けたことがあります。それなら「金 仮説」とか書くでしょうに。

この帰無仮説、先ほども説明しましたが 「差がない」という仮説です。自分が知りた いことや自分の立てた仮説が「差がある」 だとしても、帰無仮説は「差がない」なの です。つまり、統計的検定は皆「差がない」 ことを検証しています。それで検証できるか できないかで、差がないのかあるのか(厳 密には「差がないとはいえない」のか)を 検証しています。それだけは忘れないでく ださい。

今回の例で行くと、「カイ2乗値」の「p値」 が0.05より小さいので帰無仮説を棄却(否 定)し、一般的には「男子学生と女子学生 で虫歯の比率が同じとはいえない」と考え ます。この「いくつより小さければ帰無仮説

を棄却するかという値をαとか有意水準 といい、事前に決めておくことが必要とさ れています。

なお、厳密には帰無仮説は「差がない」 という仮説なので、仮説が棄却された場 合は「差がないとはいえない」ということ になります。ここが諸説あるところで、厳 密に「差がないとはいえない」の考えを守 る人たちと、「『差がないとはいえない』と いうのなら、『差がある』といったっていい じゃないか」という人たちがいます。確か に厳密には、差があることを検証はして いないので、「差がないとはいえない」に なるのですが、「差がある」としたほうが 何かと便利であることは確かです。統計 の本を見ても、「差がないという仮説が棄 却された」という表現のものから「有意差 あり」とするものまでいろいろです。

ただ、「差がないとはいえない」「有意差 がある」のどちらで表現する場合でも、設 定した有意確率と同じ割合だけ間違った 結果になる確率があることは知っておい てください。本当は有意差がない(帰無仮 説が棄却できない)にもかかわらず、たま たま結果が棄却できると考えられる範囲 にいってしまう確率は、有意水準と同じだ

けあるのです。それで有意水準のことを 危険率ともいいます。

なんだか教科書のように説明が長くなっ てしまいました。統計学に慣れ親しんで いる方には当たり前の話ですが、ここを初 めての人に理解してもらうにはなかなか 大変ですし、これが理解できれば統計的 検定については理解できていると判断し てもいい事象だと思います。

次回は、「分析」メニューの「分散分析」 を中心に説明をしたいと思います。

千葉大学医学部疾患プロテオミクス寄附研究部門 データベース・インフォマティクス担当

ヤンセンファーマ株式会社 研究開発本部 臨床統計部 統計解析グループ マネージャー

大橋 涉

医学博士

コラム「SAS四方山話」

第14回 医学研究のススメ(2)

研究開始に向けて

前回(春号)では、ご自身が医学研究を 開始するに当たって注意すべき点につい て、基本的な部分について述べさせてい ただきました。前回の説明でもまだ疑問 は持っているが、論文の執筆なんて自分 には無理、などと現段階では思っておら れる方が、もしもこれを読んで、将来的に はやってみようという気持ちになっていた だけるのであれば、それは本当に筆者 冥利につきます。これから「研究デビュー」 する予定の皆様のための心得と申します か、本当に基本的な部分だけですので、 これさえ満たしてさえいれば万事OKと いうわけではありません。また、医薬以外 の分野の方にとっても、新聞やニュース などで話題になる医学研究がどのように なされているのかを知っていただければ、 きっと今まで以上にニュースから見えてく るものがある…かもしれません。まずは 復習も兼ねて、ポイントを整理してみま しょう。

研究開始前に確認 (決定) すべきこと

- ①研究によって知りたいことは?仮説は? (仮説検証研究でなくとも何か知りたいことや 疑問はあるはず。これが無ければ研究は始ま らない!)
- ②新たな実験・試験を実施する必要があるか? (既存のデータやFvidenceを積み上げて仮説 の検証や、疑問を解消することは可能か?)
- ③研究の新規性は? (過去に同様の研究は存在しなかったか?)
- ④研究のデザインは? (前向きor後ろ向き、介入or観察など)
- ⑤研究の評価項目(Endpoint)は? (研究者の疑問とリンクしているか?主要評価 項目と副次評価項目は区別されているか)
- ⑥評価項目の集計・解析方法は? (解析手法、用いるソフトウェアは)
- ⑦研究の実施可能性は? (時間、予算、倫理面など)

手順リスト

論文を書く目的とは?

まずはご自身の問題意識を整理整頓して、 それらの解決方法を見出してみましょう*1)。 基本的に、研究論文として認められるた めには新規性*2)が必要であり、既に結果 が存在するものに関しては研究の実施が 認められません。仮に実施したとしてもそ れは二番煎じとして扱われるため、論文等 の投稿はまず認められません。そもそも研 究の目的は、自身の疑問や課題を明らか にすることなのですが、明らかにした結果 は「研究者だけのもの」ではありません。 その結果は論文化されることによって、初 めて説得力のある根拠(Evidence)として 社会に浸透することになるのです。最終的 にはそれらの Evidence が積み上げられる ことによって、医学の発展に寄与すること になります。その結果、多くの患者様が救 われることになるわけです。臨床医の皆様 からはよく、「診療が忙しくて研究している ヒマが無い」「研究論文の書き方が分から ない」などというお話を聞くのですが、筆者 としてはそれでも敢えて書くことをお勧めし たい、いや、むしろ書いて欲しいと思ってお ります。皆様の貴重なご経験をご自身だけ のものにしてしまうのは非常に勿体無く、 もしかしたら他に救われる患者様がいらっ しゃるかも知れないと思うからに他なりま せん。書き方が分からないなどという理由 で、皆様の貴重なご経験がそのまま埋も れてしまうのは、本当に勿体無いことなの です。



研究方法と説得力

次に研究の方法としては、大きく右表の 2通りが考えられます。各々の研究者の都 合や事情(時間的、予算、実現可能性な ど)と相談しつつ、研究の方針を決定する 必要があります。時間や予算を無限に持っ ている研究者はまず存在しませんので、 実は可能な研究方法はこの段階で限られ てきます。たとえば臨床試験を実施したく とも、そのためには莫大な予算と時間を 要しますので、残念ながら研究者個人で 実施することはほぼ不可能となります。 研究計画を立案して科研費を取得するか、 製薬企業などのスポンサーと共に進める かですが、いずれにしても厳しいものがあ ります。

	既存のデータや文献などを吟味	新規の臨床試験や実験を行う
時間	短時間での解決が可能	短くても数カ月、長ければ数年レベルの時間を要する
コスト	文献や資料代が必要だが、臨床研究ほど のコストは要しない	非常にコストが高く、大量の研究費用を要するため個人では実施不可能
	・あまり高いものは望めず、上位ジャーナル への投稿が厳しい場合がある	・高くなりやすく、結果次第ではいわゆる上位 ジャーナルへの投稿が可能になる
説得力	・基本的に後ろ向き研究であり、後付解析に なるため、説得力は低めになる場合が多い	・一つの試験からいくつもの結果(論文)を導 くことが可能であり、社会及び研究者の実績 への貢献度は高い
	・時に新規性が認められないような場合も ある	・新規性が認められないような場合はほとんどない
その他	メタアナリシスを行った場合には、説得力が高くなる場合もある	IRB(治験審査委員会:Institutional Review Board)の了承やヘルシンキ宣言への配慮、さ らにはICMJEの公認サイトであるU-MIN (University Medical Information Network) 等への臨床試験登録が必要になる

表1: どちらの研究スタイルを選ぶか(説得力やコストと相談)

メタアナリシスとは?

デザインや規模の異なる複数の臨床試 験のデータを統合して解析すること。一つ 一つの試験では見えてこなかった、サブグ ループによる解析が可能になり、特にRCT のメタアナリシスは、Evidence Levelが最 も高いと言われている。各試験の症例数や 平均値、標準偏差等の基本統計量のみを 用いて行う場合と、ローデータ(生データ)を 用いて行う場合があるが、後者の信頼性の 方が高い。

一方では、過去に実施された臨床試験の データを収集するため、「研究者にとって都 合の悪い試験は取り入れない」などのバイ アスの危険性も指摘されている。それゆ え、最もEvidence Levelが高いとされる分 類に対し異を唱える人々もいる。

(メタアナリシスは両論あるのでさておき ですが) やはり実際に臨床試験を行った 場合の説得力の方があると言えそうです。 研究者としては、自身の事情が許す中で 最大限のEvidenceを収集して、可能な限 り説得力の高い論文を発表したいと考え るのです。それらの説得力は、一般的には IF (Impact Factor*3) という指標により 示されるわけですが、この値は高いほど 社会的に影響力のある雑誌であると言え ます。論文は研究者の実績となり、時に学 位の授与や国からの研究費の獲得等に おいて非常に重要な位置づけとなるという 側面もございます。研究者ができるだけIF の高い雑誌に掲載されようと一生懸命に なるのは、このような現実的な側面も否定 できません。ですが、研究費の獲得はさら にEvidenceの精度を高めるために必要 なことでもあります。そのためには実績が 求められるのも、また事実なのです。IFの 高さ=論文の価値ではありませんが、それ らを抜きには成り立たない部分もあること は確かなのです。

Evidence Level 研究デザイン

レベル 研究のデザイン Randomized Clinical Trial (RCT)のメタアナリシス RCT(ランダム化比較試験) Ιb デザインされた比較研究 Πa (非ランダム化) Пh デザインされた準実験的研究 デザインされた非実験的記述研究 Ш (比較研究・相関・症例研究等) 低 専門家・権威者の意見・経験

表2: Evidence Levelと研究デザイン (Oxford Centre for Evidence-based Medicineの分類(2001)より作成)

上記は研究の種類とEvidence Levelにつ いてまとめられたものです。メタアナリシス についてはさまざまな意見があり、 Evidence Levelが一番であるという意見 に対しては異を唱える人々も少なくありま せんが、少なくとも Ib ~IVの順番に関して は、筆者も特に異を唱える必要はないと 考えます。たとえば表2の右側の研究デザ インの定義は、実際に実験・試験を行う わけではありませんので、Evidence Level としてはⅡb~Ⅲ程度と、臨床試験を行っ た場合よりは Evidence Level が低くなりま す。特に記述的研究の場合には、新規性 など基本的な部分で認められなかったり

> する場合もありますので、研究の 開始前にあらゆるオンラインの データベース (Pubmed や Medline、 日本語であれば医中誌など*4)等 で念入りに類似の研究を調べたり 専門家へ相談に行ったりする必要 があります。

> 特に注意点として、新規性は各 雑誌の編集者(Editor)や査読 者(Reviewer)の判断による部分 も大きいですので、必ずしも明確 な基準があるわけではございま せん。あるジャーナルが「新規性 なし」と却下(Reject)した論文で も、別のジャーナルでは採択 (Accept) されることはあります

ので、この部分については(特に研究に慣 れていない段階では)、ご自身の判断で決 めつけてしまわないことをお勧めします。

Critical Readingのススメ

冒頭にある手順リストの23にも示させ ていただきましたが、実は研究の立案から 開始までの間には、研究者はそれなりの 論文数を読み込まなければなりません。 これは研究を行う場合にはどうしても避け て通ることのできない、言うなれば通過儀 礼のようなものです。しかもIFの付与され ている論文は全て英文ですので、それなり の語学力も必要になってきます。記載して あることを理解できるだけでなく、時には その内容の真偽について問わなければな らない場合もありますので、当然語学力だ けでどうにかなるというものでもありませ ん。内容の真偽などと申しますと、まさか 「ウソを書いている論文があるのか?」な どと言われてしまいそうですが、論文の著 者が意図的にウソを書いてくることはまず ありません。ウソと申しますと、たとえば 「行ってもいない試験をいかにも行ったか のように書いた」「データを捏造した」など の行為を連想される方もいらっしゃるかも しれませんが、その手のお話は近年では、 まずお目にかかれることはほとんどありま せん。この場合はウソと言ってしまうと少々 語弊があるかもしれません。ウソと申しまし ても、それは研究者の悪意によるものでは なく、研究者のミスや不注意(時には無 知)によるものがこれに該当します。たと えば前号で説明させていただきました、 背景因子を揃えないままのχ²検定の単純 比較による結果などはその典型例です。

論文には確かに査読者がおりますし、 大半の場合はチェックを受けるのですが、 実は…時に査読をすり抜けてしまうことも あります。それ以上に、当時の査読者が 気付かなかった問題点を新たに指摘する ことにより、論文に新たな視点を与える ことで、論文はさらなる価値を見出すこと が可能になるのです。出版された論文を 単に受け入れるのではなく、そのような意 味 でも、「論 文 の 批 判 的 吟 味 (Critical Reading)」は極めて重要な作業であると 言えます。批判的と言うと、何故かあら探し であるとか、とかく否定的なものととらわれ てしまいますが、実はこれこそが本来の論 文の読み方なのです。ご自身が Reviewer になったつもりで読むことにより、自然とご 自身の執筆時の基本と申しますか、気を つけるべき部分なども見えてくるのです。

最後に、論文の基本的構造としてICMJE は「IMRAD構造」を推奨しています。これ は、「Introduction, Method, Result and Discussion | の略で、それぞれ論文 の主要な構成要素を示しております。

これ以外にもAcknowledgement (謝辞) や Reference (参考文献)、Conclusion (結論) などがございますが、基本的にはIMRAD に従っていると思って間違いありません*⁵⁾。 IMRAD別に、筆者なりの着目点を記載さ せていただきます。

今号はSASの話がありませんでしたが、 次号では各論になりますので、もう少し SASの話ができると思います。

	着目すべき点
Introduction (緒言)	研究を行うに至った経緯や背景などが記載されているか。この研究の目的 や仮説、新規性などが明確に記載されているか。
Method (方法)	研究の対象は何か。特に研究の対象となった集団についてのプロフィール (年齢、性別、既往歴など)が明確に記載されているか。その集団が選択された理由および、別の集団が選択されていない理由は何か。主要評価項目、副次評価項目はそれぞれ何であり、仮説や問題意識とはリンクしているか。この方法をどうして選択したのか。統計処理を行った場合はその手法が適切なものであるか。用いたソフトウェア名は記載されているか。介入・観察・割付の方法はどうか(ランダム化、盲献化などを行っている場合は、その方法や遵守の度合いなどを示しているか)。利用したデータの出所は明確であり、信頼できるものか。
Result (結果)	仮説や研究者の疑問は解決できているか。統計処理の結果は記載されているか。後付の解析になっていないか。解析対象となった集団の数は明確に記載されているか。アスタリスク以外に、p値は記載されているか。
Discussion (考察)	他の研究や先行研究と比較して、当該研究による新しい知見は何か明確に記載されているか。その重要性について記載されているか。研究結果からの極端な飛躍(たとえば有意でないp値に対し「傾向が見られる」などの記述)はないか。

表3: IMRAD別の着目点

- *1) そもそもこれがなければ研究そのものが成立しないのですが、時に学位などのために、嫌でも問題意識を ひねり出して論文を執筆しなければならない場合もあります(本来は何らかの問題意識があるからこそ 進学するものなのですが…)。
- *2) 国際医学雑誌編集者委員会(International Committee of Medical Journal Editors) の定める「牛物医 学雑誌への統一投稿規程(2010年改定)」にも、「抄録はその研究もしくは観察における新規かつ重要な 特徴が強調されていなければならない」と明記されています。そのぐらい新規性は大事なのです。
- *3) 一言でいえば「その雑誌の影響度」です。(昨年および一昨年に掲載された論文が本年に引用された回数) /(昨年の掲載論文数+一昨年の掲載論文数)で示され、この値が大きいほど影響力があると言われてい ます。筆者個人的には5点以上でもすごいと思っている次第ですが、NatureやNew England journal of Medicineは30点以上です。多くの英文誌には付与されておりますが、英文誌でも付与されていないものも 多数あります。
- *4) 全てオンラインの医学論文のデータベースです。有名なのはPubmedですが、そのほかにはMedline、日本 語ですと医中誌があります。
- *5) 雑誌により異なりますが、各雑誌には投稿規程がございます。基本的にはIMRADに従っているのですが、 優先されるのは各雑誌の投稿規程です。

Q & A



Q&A

- ショートカットキーにマクロプログラムを登録する方法
- デンドログラム、樹形図の表示に関して
- ライヤンス 更新をバッチで行う方法
- BIクライアントプロダクトの機能制限について
- 変数名をラベルに割り当てる
- Windows 7にてJREを指定することができない
- ▶ 存在しないクラス変数の値をMEANSプロシジャの出力に 含める方法





頻繁に実行するマクロプログラムをショートカットキーに 登録して、ワンクリックで呼び出すことは可能ですか?



F9キーでショートカットキーウィンドウを起動した後、任 意のショートカットキーを次のように定義することでマクロを呼び出せるようになります。

submit '%test';

なお、マクロプログラム(%test)は事前にコンパイルしてください。



TREE プロシジャにてデンドログラム、樹形図を表示しています。一つのグラフではなく、複数のグループに分けて表示させることはできますか。



最初に、PROC TREE ステートメントにおけるHEIGHT= オプション、LEVEL=オプションを用いて、複数のグルー プ分岐を示すグラフを作成します。

例

PROC TREE DATA=treedata HEIGHT=NCL LEVEL=5; RUN;

ここでは、クラスタの数を基準(HEIGHT=NCL)とし、グループに分岐する基準値を5(LEVEL=5)としていますので、5つのグループ分岐となります。

作成されたグラフでは、クラスタ番号を示す 'CL5' などの表示があります。再度、TREEプロシジャを実行する際、ROOT=オプションにてクラスタ番号を指定することによって、各グループごとにグラフを作成することができます。

例

PROC TREE DATA=treedata ROOT='CL5';
RUN;



多数のマシンに対してSASライセンスの更新作業を行う場合、ライセンス更新ツールをバッチ実行することはできますか?



SAS ライセンス更新ツールはバッチ実行が可能です。 コマンドプロンプト、もしくはバッチファイルに次のコマンドを 記述して、管理者権限を持ったユーザで実行してください。

SAS インストールフォルダ \core\sasinst\sasrenew\sasrenew.exe" -s "datafile: c:\SAS92_xxxx_xxxxx.txt"

例 SAS® 9.1.3 の場合

"C:\Program Files\SAS\SAS 9.1\core\sasinst\sasrenew \text{\fome}\sasrenew.exe\" -s \"datafile: c:\SAS\91_XXXX_\Win_\Wrkstn.txt\"

例 SAS® 9.2 の場合

"C:\Program Files\SAS\SasFoundation\9.2\Core\sainst \Sasrenew\sasrenew.exe" -s "datafile: c:\SAS\92_XXXX_\text{Win Wrkstn.txt"}

なお、更新処理のログは以下のファイルに出力されます。

SAS インストールフォルダ ¥setinit.log

例 SAS 9.1.3 の場合

"C:\Program Files\SAS\SAS 9.1\Setinit.log

例 SAS 9.2 の場合

C:\Program Files\SAS\SASFoundation\9.2\setinit.log



ユーザーによってSAS® Enterprise Guide 4.3で行う 作業の分担を考えています。特定のユーザーは特定の 機能のみ使用できるような機能はありますか?



SAS® 9.2より、メタデータに登録されているユーザーに対して、BIクライアントプロダクトの機能を割り振ることが可能となる「役割」という機能が追加されました。

データの加工のみ有効とする役割、データの分析機能のみ有効とする役割を作成し、任意のユーザーに対してそれらを割り当てることで特定のユーザーに対して使用可能な機能を割り当てることが可能となります。

例えば、ユーザー A がメタデータ上に登録されており、Enterprise Guide 4.3 にある「データ」機能のみ使用可能にする場合、設定方法は次のようになります。

- 1. SAS管理コンソールを起動し、メタデータを編集可能なユーザーで接続します。
- 左ペインの「プラグイン」タブを選択し、
 「SAS管理コンソール」→「環境管理」→「ユーザーマネージャ」
 を右クリックします。
- 3. 「新規作成」→「役割」を選択します。
- 4. 「新規役割のプロパティ」画面の各タブの設定を次のように 設定しOKボタンを押します。

「一般」: データ機能のみ「メンバ」: ユーザー A

「機能」: ツリーより、「アプリケーション」→「Enterprise Guide 4.3」

→「データ」以下の項目にチェック

※通常、SASを使用するユーザーは暗黙的にPublic グループとSASUSERS グループに属しています。このため、任意のユーザーに役割を割り当てた場合でも、それらのグループに割り当てられている役割の影響を受ける可能性があります。この場合、それらの役割を外すことでユーザー Aに設定した役割を有効にすることが可能です。

[Usage Note 40599: Assigning users to a custom role that has limited capabilities might still have access to additional capabilities even after applying the custom role] http://support.sas.com/kb/40599

役割の詳細に関しましては、次のドキュメントをご参考にしてください。

[SAS 9.2 Intelligence Platform: Security Administration Guide: Role Definitions]

http://support.sas.com/documentation/cdl/en/bisecag/611 33/HTML/default/viewer.htm#a003142396.htm



変数名をラベルに割り当てたいです。どのような方法が ありますか。



CONTENTSプロシジャを使用することで、変数名、順序などのディスクリプタ部分を参照、データセットに出力することができます。変数名をマクロ変数に格納することで

実現可能です。

例えば、下記のような例が考えられます。

例

```
DATA sample:
  a=123;
  b=234:
RUN:
%MACRO setlabel(srcdat);
PROC CONTENTS DATA=&srcdat OUT=_tmp(KEEP=name) NOPRINT;
RUN:
DATA _NULL_;
  SET _tmp;
  CALL SYMPUT("n",_n_);
  CALL SYMPUT("var" | |left(_n_), name);
RUN:
DATA &srcdat;
  SET &srcdat;
  LABEL
  %DO i = 1 %to &n;
  &&var&i = "&&var&i"
  %END:
RUN:
%MEND setlabel;
%setlabel(sample);
```

Q & A



Windows 7にてSAS® 9.2をインストールしているのですが、JREの指定の箇所でJREを指定することができません。どうしたらよいでしょうか。



Windows 7ではインストール画面にてJREが自動で選択されません。JRE 1.6.0_14をダウンロードし、以下の手順にてインストールしておく必要がございます。

1. 以下のホームページにアクセスします。

https://cds.sun.com/is-bin/INTERSHOP.enfinity/WFS/CDS-CDS_Developer-Site/en_US/-/USD/ViewProductDetail-Start?ProductRef=jre-6u14-oth-JPR@CDS-CDS_Developer

- 2. 「Platform」欄より「Windows」を選択し、以下のチェックボックス にチェックを入れ、「Continue」ボタンを押します。 [I agree to the Java SE Runtime Environment 6u14 with JavaFX 1 License Agreement.]
- 3. 「jre-6u14-windows-i586.exe」をクリックします。
- 4. 「セキュリティの警告」ダイアログボックスが2回表示されますので、2回とも「はい」を選択します。
- 5. 「ファイルのダウンロード」ボックスが表示されますので、「保存」 ボタンを押下し、ファイルを保存します。
- 6. ダウンロードしたファイルをダブルクリックし、ウィザードの指示 に従ってインストールを実行します。
- 7. JREをインストール後、SAS 9.2をインストールします。
- 8. インストールの途中にて既存のJREのインストールを聞かれる ダイアログボックスが表示されます。ここで「参照」ボタンをクリッ クし、JRE 1.6.0_14へのパスを選択します。通常は以下のフォル ダにインストールされますので、以下のフォルダを選択します。

64ビットWindows:

C:\Program Files(x86)\Java\Jre6

32ビットWindows:

C:\footnote{C:\foo

9. その後 SAS のインストールを続行します。 関連する SAS Note につきましては以下の URL に掲載されて おります。

[Installation Note 40501: SAS 9.2 requires JRE 1.6.0_14 32-bit on Microsoft Windows 7 system and JRE 1.6.0_23 32-bit on Microsoft Windows Server 2008 R2 system]

http://support.sas.com/kb/40/501.html



MEANSプロシジャで要約統計量を出力する際、存在しないクラス変数の値を表に含めることはできますか? 具体的には、データには1月から6月のデータのみが存在

している状態で、出力は1月から12月まで出力したいと考えています。



ユーザ定義フォーマット作成後、MEANSプロシジャにて、COMPLETETYPESオプションおよびCLASSステートメントのPRELOADFMTオプションを使用することで、

意図した出力が生成されます。

例

```
PROC FORMAT;
                        /* ユーザ定義フォーマット作成 */
 VALUE mfmt.
       1='1月' 2='2月' 3='3月'
        4='4月'5='5月'6='6月'
       7='7月'8='8月'9='9月'
        10='10月' 11='11月' 12='12月';
RUN:
DATA sample;
  DO month = 1 TO 6;
   DO i = 1 TO 30;
     val = 1000+(1000*ranuni(543));
     OUTPUT:
   END:
  END;
 FORMAT month mfmt.;
PROC MEANS DATA=sample SUM MEAN NONOBS;
  CLASS month;
  VAR val;
RUN:
                       /* completetypes オプション */
PROC MEANS DATA=sample SUM MEAN NONOBS
                      COMPLETETYPES:
                         /* preloadfmt オプション */
  CLASS month / preloadfmt;
 VAR val;
RUN;
```

オプション無

month	合計	平均
1月	45389.76	1512.99
2 月	45379.07	1512.64
3 月	41152.30	1371.74
4 月	44917.96	1497.27
5 月	46752.22	1558.41
6 月	46306.09	1543.54

オプション有

month	合計	平均
 1 月	45389.76	1512.99
2 月	45379.07	1512.64
3 月	41152.30	1371.74
4 月	44917.96	1497.27
5 月	46752.22	1558.41
6 月	46306.09	1543.54
7月	•	•
8月	•	•
9 月		
10 月	•	
11月		•
12 月		

のお知らせ

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

●「SASによる症例数設計と中間解析」コース(2日間)

2011年9月27日(火)~28日(水) 10:00 ~ 17:00 (東京会場)

147,000円(税込)/※チケットのお取り扱いはありません

- ※1. 過去数回にわたって開催した「POWERとGLMPOWERプロ シジャによる症例数設計」コースの内容と今回の症例数設計の 内容については同一となりますのでご注意ください。
- ※2. 本コースは、英文テキストを使用し日本語で説明を行います。 また、補助資料を配付します。

「受講対象]

製薬企業等で医薬研究のデザイン、解析業務に関わる方 [学習内容]

POWERとGLMPOWERプロシジャを用いた症例数設計、中間解析 のSEQDESIGNSプロシジャによるデザイン、および SEQTESTプロ シジャを用いた解析の詳細について解説と演習を行います。 SAS® 9.2で追加されたSEQDESIGNプロシジャとSEQTESTプロシ

[Web Page]

ジャを習得できます。

www.sas.com/jp/training/course/nocd_ia.html

●「医学研究における経時データ解析」

[日

2011年10月24日(月)~25日(火) 10:00 ~ 17:00 (東京会場)

147,000円(税込)/※チケットのお取り扱いはありません。

「受講対象]

SASによる統計解析に従事する方

[学習内容]

医薬統計専門家により以下の内容で行います。

多くの医学研究においては、研究対象者毎に時間を追って反応変 数を繰り返し測定することがあり、このような研究は経時観察研究 などと呼ばれます。

その研究目的の一つは、反応変数と時間、治療効果あるいは他の 説明変数との関係を定量的に評価することです。同一対象者内の 反応変数には相関が生じることから、統計的に妥当な推論を行う ためには測定単位に起因する相関を考慮する必要があります。そ のような反応変数間の相関を考慮する方法として、線型混合モデ ル及び一般化推定方程式(GEE)による解析方法が近年注目を浴 びてきています。このコースでは、線型混合モデルと一般化推定方 程式の理論的背景について講義し、実際の医学データの解析を 通してMIXEDプロシジャ及びGENMODプロシジャの使用法につい て学びます。

- ・線型混合モデル及び一般化推定方程式(GEE)に関する基本概念
- MIXEDプロシジャ及びGENMODプロシジャの使用法
- ・医学分野における実データの解析

[Web Page]

www.sas.com/jp/training/course/longdata_analysis.html

新規トレーニング・コースの開催のご案内

●「SASによるベイズ解析」コース (2日間)

2011年9月1日(木)~2日(金) 10:00 ~ 17:00 (東京会場)

格]

115,500円(税込)/チケット捺印数2

「受講対象]

ベイズ分析によるアプローチに興味をお持ちの、生物統計、疫学、 社会科学を専門あるいは仕事に従事している方

[学習内容]

本コースは、PHREG、GENMOD、MCMCプロシジャを使用した、ベ イズ分析に焦点をあてています。多くの事例は、臨床試験の分野か らご紹介します。

- •ベイズ分析の概要
- ・PROC PHREG、PROC GENMODによるベイズ分析の解説
- ・ベイズ分析による事前分布の利用
- ・PROC MCMCを使用した、臨床試験のデザインにおけるベイズ 分析アプローチの解説

[Web Page]

www.sas.com/jp/training/course/sas_bayesian.html

●「SAS Financial Management:必須要素」 コース (1日間)

程]

2011年9月21日(水) 10:00 ~ 17:00 (東京会場)

57,750円(税込)/チケット捺印数1

[受講対象]

予実管理、連結決算、レポート作成、分析担当者や予算編成と業 務計画(販売計画など)の作成責任者

[学習内容]

本コースはSAS® Financial ManagementとSAS Financial Management のワークスペースとディメンジョンの概要を解説します。

[Web Page]

www.sas.com/jp/training/course/fm1.html

● [SAS Financial Management:レポーティング] コース (1日間)

程]

2011年9月22日(木) 10:00 ~ 17:00 (東京会場)

57,750円(税込)/チケット捺印数1

[受講対象]

予実管理、連結決算、レポート作成、財務分析担当者 [学習内容]

本コースはSAS Financial Managementの管理者ツールの概要を解説 します。受講者は会計レポートの作成を行います。リアルタイムの情報 とオンデマンドの連結会計を使用した、結果の分析方法も学びます。 [Web Page]

www.sas.com/jp/training/course/fm_reporting.html



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/jp/training/

その他、トレーニングに関する情報については、右記の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。

トレーニング担当

T E L: 03-6434-3690

F A X: 03-6434-3691

E-mail: JPNTraining@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

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

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



SAS Technical News入手

SAS Technical Newsは、右記のURLから入手できます。

http://www.sas.com/jp/periodicals/technews/index.html

発行:SAS Institute Japan株式会社



SUMMER 2011

■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:03-6434-3680 FAX:03-6434-3681



POWER

SAS Institute Japan株式会社

本社 〒106-6111 東京都港区六本木6-10-1 六本木ヒルズ森タワー 11F Tel 03(6434)3000 Fax 03(6434)3001

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

www.sas.com/jp