

# SPSS —汎用統計パッケージ—

国立公衆衛生院\* 林 正 幸

(1996年7月 受付)

## 1. はじめに

電子計算機、とりわけ1960年代から1970年代にかけての20年間に主流であった大型汎用機は、特別な技術をもったオペレーターと特別な知識を必要とするプログラミングが可能な利用者が存在して初めて稼働するものであり、加えて非常に高額なことから、その利用はごく限られた施設と目的の範囲でしかなかった。しかしながら、1980年代に入り比較的安価な個人用コンピュータ(PC)と操作が容易なOS(Operation System)が実用の段階に突入し、便利で優秀なソフトウェアの開発と相まって飛躍的に発展を遂げ、長く汎用機の世界であった統計解析やシミュレーションも手軽に手元で実行できる状況となった。1990年初頭にはOSにUNIXを利用するEWSも一般で用いられるようになり、同時にPCのOS(MacやMS-DOS)も飛躍的に発展した。PCのOS上の問題点であった利用可能なユーザーメモリーの制限もWindows環境が提供されるにいたり解決され、統計解析に関する限りよほど巨大で大量のデータを取り扱う必要性がない限り汎用機やEWSに依る必然性はなく、現在ではほとんどのデータ処理は個人でも持つことのできるPCにWindows OSを載せた環境で十分可能な状況となっている。

統計解析ソフトも電子計算機の性能向上と共に発展し、当初はデータの総量に制限があったり、大量のデータは扱えるが操作が複雑で、解析手法も一般的なもののしか用意されておらず、特殊な解析は自分でソフトウェアを開発して利用しなければならないものであった。大型汎用機の時代から実用性や統計学的正確性を兼ね備えた統計解析用パッケージソフトとしてSASやSPSS, Stat Viewが著名で業務や研究に良く利用されている。SASやSPSSは統計の基礎のあるユーザーにとって便利に利用できるよう工夫されているため、統計の初心者にはモデルの適応などの間違いを冒す可能性が高い。本来、統計解析を進める前に自己の持つデータの性質を充分知り、その上で統計解析に取りかかることは当然である。統計を道具として使うユーザーをこれらの統計ソフトの利用者と考え、ソフトの側の問題とは考えにくく、ユーザーの側において事前の吟味の必然性を喚起する以外解決法はないと考える。今回、統計ソフトを利用しようとするものでコンピュータの操作に不慣れな利用者にとって、ユーザーインターフェイスが良好であり、利用に際して特別な契約条件もなく、しかもコスト・パフォーマンスにも優れていると考えられるSPSS/Windows 95 (SPSS Ver. 7.0以後)を取り上げ、その特徴を解説する(尚、1997年3月現在、SPSS Windows 95版 Ver. 7.5J暫定版が販売されている)。

## 2. SPSSの沿革

1968年初頭、日本・アメリカ・インド・ナイジェリアの4ヶ国によって、ある分野における国際的な協同研究が行われていた。その中で、協同の企画による共通の調査を実施し、その結

---

\* 保健統計人口学部：〒108 東京都港区白金台4-6-1。

表1. SPSS 開発の経緯.

年 月	事 柄
1968年 1月	SPSSのプロトタイプ開発される
1969年 5月	SPSS 国内上陸
1973年 4月	KUSPSS V3.0 京都大学で稼働。KUSPSSのKUは、Kyoto Univ.
1975年 4月	SPSS Inc. (米国シカゴ) 非営利団体として設立
1976年 4月	KUSPSS V6.0 京都大学、北海道大学、名古屋大学で稼働 この頃より、国内でSPSSガイドブックが刊行され始める 「SPSS統計パッケージ」(三宅・山本(1976)、三宅 他(1977))、データ解析入門 SPSSへの招待(司馬(1977))
1978年 4月	SPSS Inc. (米国シカゴ) 営利団体として設立
1979年11月	日立(VOS2/VOS3)版 SPSS 東京大学で稼働
1980年 4月	K-SPSS V6.0 筑波大学、ACOS6に試験的に移植。K-SPSSのKは移植の基となったカンザス大学版の意味
1981年 4月	SPSS V8.0(ACOS4) 青山学院大学で稼働
1981年12月	SPSS V8.0(ACOS6) 東北大学、大阪大学で稼働
1986年 1月	「新版SPSS x I. 基礎編」(三宅・山本(1986)) 発刊される
1988年10月	日本法人設立。英語版製品の販売と日本語版製品の開発と販売
1988年11月	SPSS/PC+ V3.0 DOS英語版出荷 その他、大型・中型汎用機版、EWS版等の製品供給を開始
1989年 1月	「やさしいデータ解析 - SPSS x による -」(司馬(1989)) 発刊
1989年 7月	SPSS/PC+ V3.0J (NEC PC9801 日本語版) 出荷
1990年 9月	SPSS R4.0 (the Macintosh) 出荷 「新版SPSS x II 解析編1」(垂水 他(1990))、 パソコン用SPSSのガイドブック「パソコンSPSS」(山本他(1991)) 「新版SPSS x III 解析編2」(三宅 他(1991)) など相次いで発刊
1995年 3月	SPSS 6.1J (the Macintosh)
1995年 6月	SPSS 6.1J (Windows)
1995年10月	SPSS 6.1J (the Macintosh (PowerMac))
1995年12月	SPSS 6.1J (Windows 95)
1996年 3月	SPSS 6.1 (UNIX(Motif))
1996年 5月	SPSS 7.0 (Windows 95/英語版)
1996年末	SPSS 7.5J 提供予定

果を各国それぞれの問題と関心に従って、それぞれの角度から国際的な比較分析を行う、といった分析や研究の自立を実現するための統計コンピュータープログラムの開発が、元米国スタンフォード大学（現米国ハーバード大学）の Sidney Verva 教授により提案され、約2年間にわたり提案に沿った統計分析用プログラムの開発が行われた。

分析用プログラムの正式名称は、“Statistical Package for Social Sciences”（SPSS：社会科学のための統計解析ソフトウェアパッケージ）と名付けられ、日本チームの開発メンバーとして、元東京大学池内一教授（故人）と元京都大学（現大阪国際大学）三宅一郎教授が参加した。（三宅（1973），SPSS Inc.（1995））今日統計解析ソフトとして国際的評価を得て利用されている SPSS は、はっきりとした利用目的すなわち現在では当然となっているオブジェクト指向をもって誕生したといえる。

その後、シカゴ大学で改良されデータ解析（統計解析）ソフトウェアとして、世界各国の教育機関や研究者に紹介され、以後、社会科学の分野のみならず、統計を利用する教育機関・研究機関で高い評価を得てきた。1995年現在世界90カ国・10,000サイト、1,500,000以上のユーザーにおいて利用されている。

わが国における SPSS は、1969年に IBM 360 シリーズで稼働した SPSS for MVS が始まりである。その後、1973年より京都大学にて KU (Kyoto University) SPSS V3.0 (FACOM 230-60)

表2. 初期の大型汎用計算機への SPSS 導入（新村（1994））。

稼働年	名 称	稼 働 先
1973	KUSPSS V3.0 for MONITOR5 on FACOM230-60	京都大学
1976	KUSPSS V6.0 for MONITOR5 on FACOM230-75	京都大学・北海道大学・名古屋大学
1979	SPSS 01-00 for VOS2/VOS3 on HITACHI	東京大学
1980	K-SPSS 6.0 for ACOS6 on NEC ACOS	筑波大学（試験的移植）
1981	SPSS V8.1 for ACOS4 on NEC ACOS	青山学院大学
1981	SPSS V8.0 for ACOS6 on NEC ACOS	東北大学・大阪大学

※ この表は、SPSS、日立製作所、NEC等から提供された資料によるが、古い資料滅失のため初期の記述が正確かどうかは不明である。

が稼働し、以後「京大版 SPSS (KUSPSS)」として普及していった (三宅・山本 (1976))。

京大版 SPSS の稼働以降、富士通・日本電気・日立製作所に代表される各国産汎用機メーカーも SPSS の供給を開始し、1976 年には KUSPSS V6.0 (FACOM 230-75) が京都大学、北海道大学、名古屋大学で、1979 年には日立製作所より VOS2/VOS3 版 SPSS 01-00 が東京大学で稼働、1980 年には、米国カンザス大学に設置されていたワードマシン (GCOS-OS) 版 SPSS を基に移植した K-SPSS V6.0 (ACOS6 NEC/ACOS, 頭の K は KANSAS 大学版の略) が筑波大学で、1981 年には、SPSS V8.0 (ACOS4 NEC/ACOS) が青山学院大学で、1981 年には、SPSS V8.0 (ACOS6 NEC/ACOS) が東北大学、大阪大学で稼働した (表 1, 表 2)。

1980 年代になると、コンピューターは小型化・高性能化が進み、ワークステーション (WS) やパーソナルコンピューター (PC) でも、以前の大型計算機と同等またはそれ以上の性能を持

表 3. 日本語 PC 用 SPSS.

発売年	製品名	OS	対応機種
1989	SPSS/PC+V3.0 J	DOS	NEC-PC98
	SPSS/PC+V3.0 J	DOS	IBM-PS/55
	SPSS/PC+V3.0 J	DOS	FMR60/70
	SPSS/PC+V3.0 J	DOS	AX
1990	SPSS/PC+V3.0 J	DOS	J3100
1991	SPSS R4.0J for the Macintosh	MAC	MAC
1993	SPSS 6.0J for Windows	Windows	Windows
1995	SPSS 6.1J for Windows	Windows	Windows
	SPSS 6.1J for the Macintosh	MAC	MAC
	SPSS 6.1J for the Power Macintosh	MAC	Power MAC
1996	SPSS 6.1J for Windows 95	Windows 95	Windows 95

つものが現れてきた。

SPSS もこのようなコンピューター環境の変化に対応し、1980 年代後半より、パーソナルコンピューター (PC) やワークステーション (WS) への対応を開始する (SPSS Inc. (1995))。

まず 1988 年に SPSS/PC+ V3.0 for DOS の IBM PC-XT, AT 版を開発し、併せて、大型・中型汎用機版、EWS 版等の改良・開発も始まった。

1989 年には、SPSS/PC+ V3.0J (NEC PC9801 版)が開発され、続けて IBM PS/55, FUJITU FMR60/70, AX, TOSHIBA J3100 用の各日本語版が提供され、日本語 PC 環境への本格的な対応を開始した (表 3)。

1990 年、SPSS R4.0 (the Macintosh) の発売が開始され、翌年には同日本語版である SPSS R4.0J (the Macintosh) が提供された。これにより DOS, Mac の 2 大 PC/OS での利用が可能となり、コンピューターの世界での汎用性が広まった。

1993 年、Windows 環境の開発とその利用者の拡大に対応し SPSS for Windows 3.1 (Ver. 6.0) の提供を開始、1995 年には Windows 版、Mac 版共にバージョン 6.1 へ移行し、GUI (Graphical User Interface) 環境にも完全対応した。

### 3. オプションプログラムの導入経緯

統計手法の多様化はコンピューターの処理能力の向上と共に望まれるようになり、SPSS においても徐々にサブプログラムが増加していった。また、性能向上のため改訂されたり、統合されたもの、改名、廃止されたものもある。中には日本で開発され日本でのみ提供された機能もあるが(京大版数量化理論や data pattern)、現在では国際的認知の立場から、IBM コンパチ大型汎用機以下 PC 版まで、これらは取り入れられていない (表 4)。

現行の SPSS 6.1J for Windows では Base に加え Professional Statistics, Advanced Statistics, Exact test, Tables, Trends, Categories, Lisrel 他いくつかのオプションが用意されており、ユーザーの目的に応じて導入できるが各々のオプションに含まれる統計手法 (SPSS Inc. (1995)) は次章に示す。先にふれた数量化理論は国内において需要が高かったため、SPSS (Ver. 6.0J) においてこの手法が使用できるよう、ユーザーサブプログラムとして希望者に頒布されたが、今後のサポートについては不明である。

### 4. 最新の SPSS Ver. 6.1 及び 7.0 の統計解析機能とその操作の実際

初期の汎用機版から 1990 年初頭までの UNIX 版、Mac 版、PC 版の SPSS はコマンドを記述して実行させるか、コマンドジェネレーターと変数ウインドウから必要なコマンドと目的の変数を選択してコマンドを記述し実行させるものであった。また、グラフ作成機能は SPSS に含まれず、後期になって他のソフトをスワップコールして実行させる機能が追加された。

1993 年 Windows 3.1 版の開発、発表に伴い SPSS 6.0J for Windows では Professional Statistics, Advanced Statistics, Tables, Trends, Categories, Lisrel を必要に応じて選択導入可能とし、基本機能であるスプレッドシート形式のデータ入力機能とグラフ作成機能が標準で Base System に装備された (図 1, 図 2)。操作はコマンドによる処理形式だけでなく、GUI (Graphical User Interface) 対応により目的のアイコンやメニューを指定する選択実行方法が提供された。従って、従来のコマンドに精通しているものは迅速に解析を進めることが可能であるし、コマンドに不慣れなユーザーは統計の基本知識があればアイコンで目的の統計処理サブプログラム (統計処理手法) を選択し、処理対象変数の指定と若干の追加オプション (付加的統計の要否や統計パラメータの指定、結果の表示法の指定やグラフ化の要否) の記述をポツ

表4. オプションプログラムの導入経緯.

年 月	事 柄
1985年	大型汎用機(IBM)版にGraphicsを追加。
1986年	大型汎用機(IBM)版にTablesを追加。
1988年	大型汎用機(IBM)版にTrendsを追加。
1989年7月	SPSS/PC+ V3.0J発売時にAdvanced Statistics, Tables, Trend を出荷。同時にMS-Chartをグラフツールとしてサポート。
1990年	大型汎用機(IBM)版にCategoriesを追加し、LisrelをVer.6から Ver.7にバージョンアップして他のオプションと同じコマンド 式に変更。
1990年4月	SPSS/PC+V3.0JにDE-IIを追加。
1990年9月	SPSS4.0J for the MacintoshではAdvanced Statistics, Tables Trendsに加えCategoriesを提供。同時にCA Cricket Graphをサ ポート。
1993年	SPSS4.0J for UNIXではTables, Trends, Categories, Graphicsに 加えOracleなどのリレーショナルデータベースからデータを読 み込むCaptureを提供。
1993年11月	SPSS6.0J for Windows (Ver.3.1) を提供。 Professional Statistics, Advanced Statistics, Tables, Trends, Categories, Lisrelを提供。 スプレッドシート形式のデータ入力機能とグラフ機能をBase Systemに装備。また、コマンドによる処理形式だけでなく、GUI 対応により簡単な指定による実行方法を提供。
1994年5月	SPSS6.0J for WindowsにCHAIDを追加。
1994年8月	SPSS6.0J for Windowsで数量化理論プログラムを配布。
1995年8月	SPSS6.0J for WindowsにExact Testsを追加。
1996年初頭	SPSS6.1J for Windows 95版をリリース。
1996年末	SPSS7.5J for Windows 95をリリース予定。

プアップするダイアログ・ボックスに対しておこなえば標準的な解析が簡単に実行できる。またアイコンとダイアログ・ボックスで選択した統計の流れを貼り付けボタンのクリックによりシンタックスウインドウに貼り付け、一連のコマンドの記述として保存し、このコマンドから同じ処理を繰り返し実行したり、コマンドの文法(シンタックス)を学習するのに便利である(SPSS Inc. (1994), Norusis (1994))。

解析対象となるデータは汎用機でよく用いられた固定長形式、PC, WS で多用される可変長形式(コンマ区切り, Tab 区切り)と呼ばれる ASCII 形式のテキストファイルのほか、エクセルや Lotus 1-2-3 の表形式データ, dBASE のデータベースシステムファイル等を直接に読み書きできる(図3, 図4)。また、第1レコードもしくは表頭に在るアルファベット8文字以内の

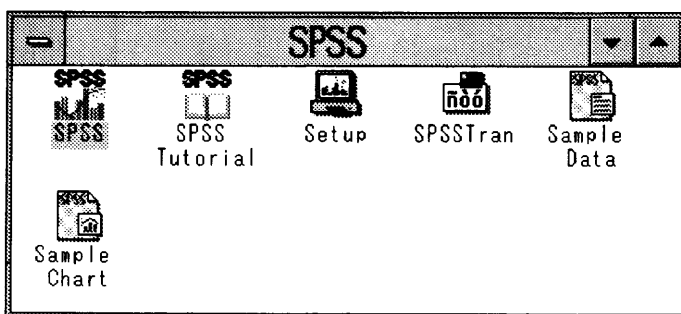


図1. SPSS for Windows (Ver. 6.1). グループアイコンとアプリケーションアイコン。

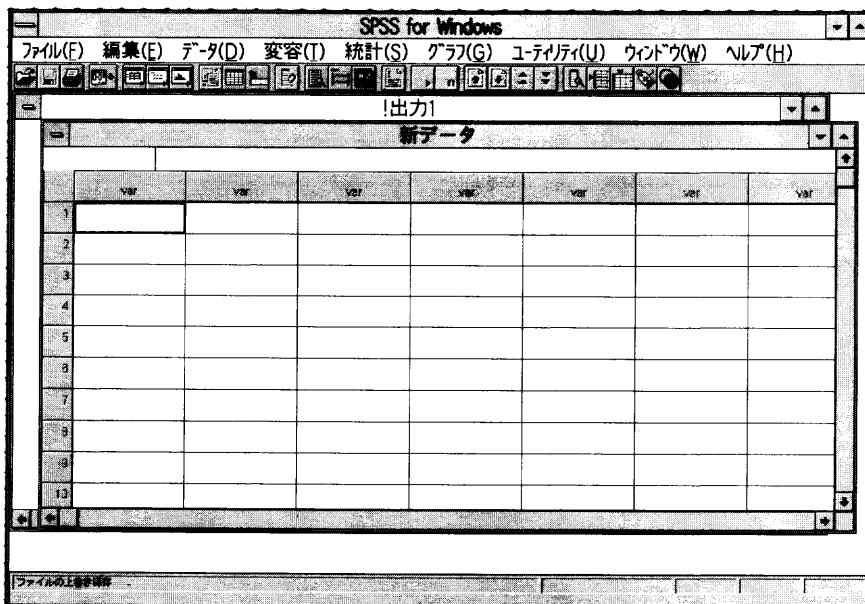


図2. SPSS for Windows (Ver. 6.1). アプリケーションウィンドウ (初期画面)。

項目名等はそのまま SPSS で用いる変数名として読み込み使用することができる。コンピュータのメモリーや補助記憶装置の物理的容量による制限と、一部の手法ごとに定められた特別な制限によるもののほか、取り扱うデータの件数と1件あたりの変数の数、すなわちレコード数と1レコードあたりの項目数の積に制限はなく、相当大量と思われるデータの解析も、時間や機器構成、費用について実用の範囲で可能である。さらに、効率的な入出力手法とファイル構造を採用しているのと同容量のエクセルや1-2-3のファイルより高速な入出力が可能となっている。また、圧縮形式でのファイル保存が用意されているので保存のためのファイルサイズを節約することも可能である。

SPSS に読み込まれたデータは SPSS のシステムファイルとしてパーソナルコンピュータシステム内に存在する(図5)。このファイルは SPSS からデータベースアイコンあるいは変容アイコンやコマンドを用いて自由に検索や加工が可能であり、加工後のデータを保存して呼び出し、再度解析することも可能である。加工は、個別のデータあるいはケース、変数毎の加除訂正はもちろんのこと、1つあるいは2つ以上の変数を用いて70種を越える組込み関数を含め

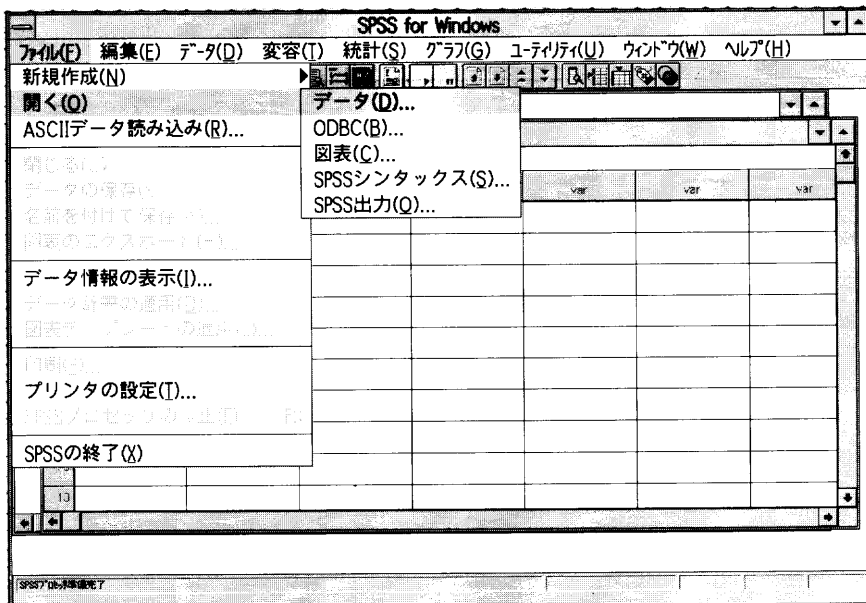


図3. ファイルメニューから、読み込みサブメニューのポップアップ。

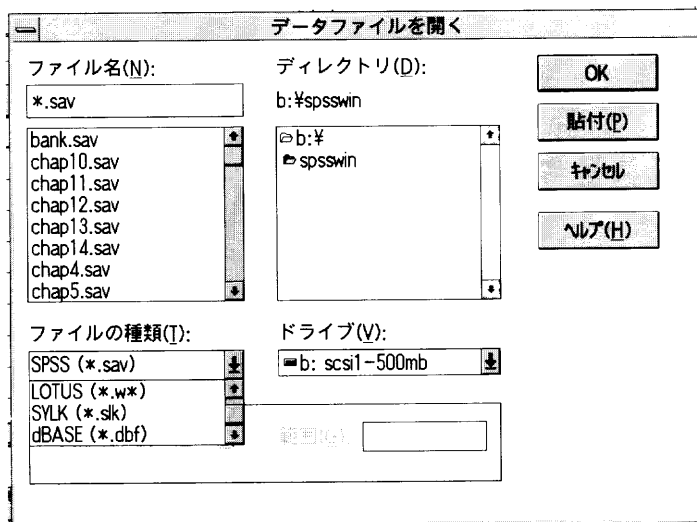


図4. データファイルのオープン・ダイアログボックス。

た演算処理や値の再割り当て機能により、元のデータを書き換えたり新しいデータとして作成、利用できる。組込み関数には、算術関数、乱数関数、統計関数、分布関数、論理関数、日付・時刻や取り出し関数、欠損値関数、ケース間関数、文字型関数他が用意されている。

データファイルは、データベース（ファイル）変容アイコンやコマンドによって他のソフトでも実行可能なソートやマージ等一般的操作のほか、論理条件を用いた解析対象の一時的あるいは恒久的選別、さらにはスプリット操作（ファイルをあるキーで並べ替えておき、男女別な



1 id	id	salbeg	sex	time	age	salnow	edlevel	work	jobcat	min
1	628	8400	0	81	28.50	16080	16	.25	4	
2	630	24000	0	73	40.33	41400	16	12.50	5	
3	632	10200	0	83	31.08	21960	15	4.08	5	
4	633	8700	0	93	31.17	19200	16	1.83	4	
5	635	17400	0	83	41.92	28350	19	13.00	5	
6	637	12996	0	80	29.50	27250	18	2.42	4	
7	641	6900	0	79	28.00	16080	15	3.17	1	
8	649	5400	0	67	28.75	14100	15	.50	1	
9	650	5040	0	96	27.42	12420	15	1.17	1	
10	652	6300	0	77	52.92	12300	12	26.42	3	

図 5. SPSS データウィンドウ。

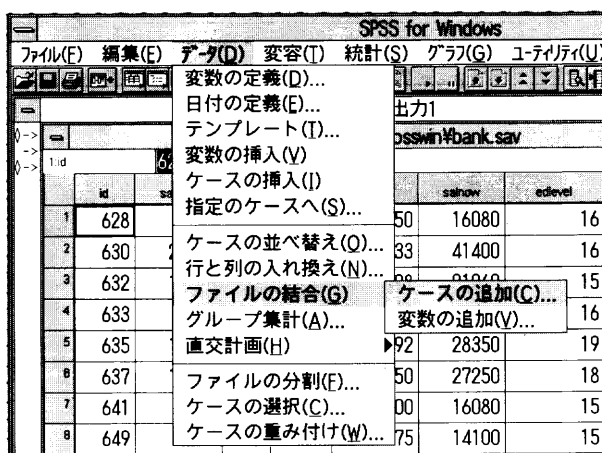


図 6. データメニューのポップアップ。

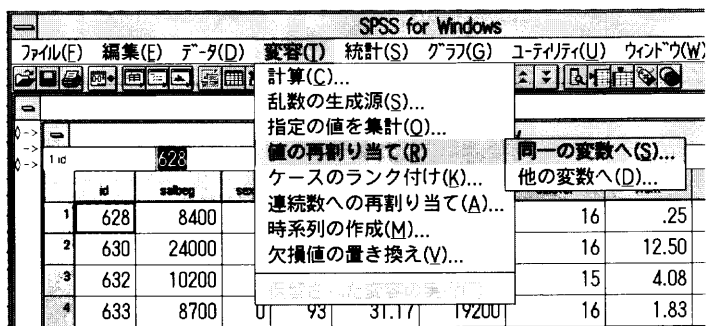


図 7. 変容メニューのポップアップ。

どそのキーの属性ごとのグループとしてファイルを分割し、それぞれのグループごとに解析を繰り返す（スプリットファイル命令では1回の統計解析指示操作で複数のグループに対して同じ手法の解析が連続して実行できる）が可能である（図6、図7、図8）。従って、ファイルや

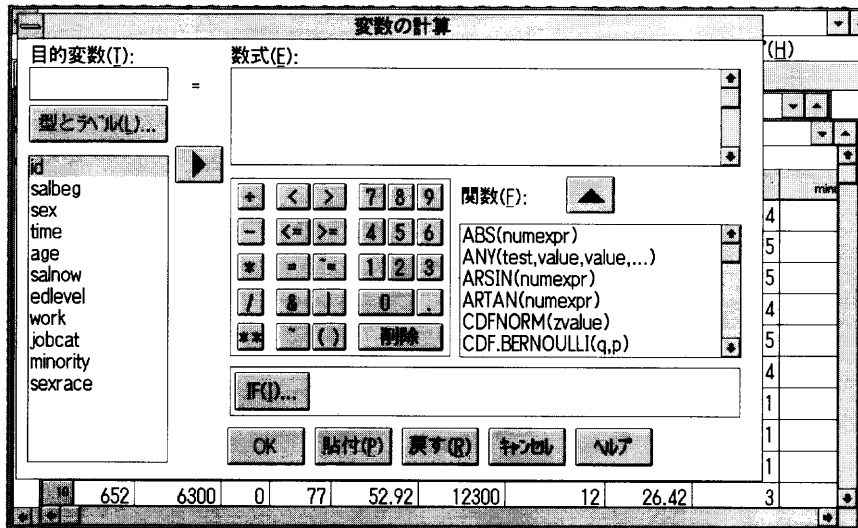


図8. 変数の計算ダイアログボックス (変容メニュー).

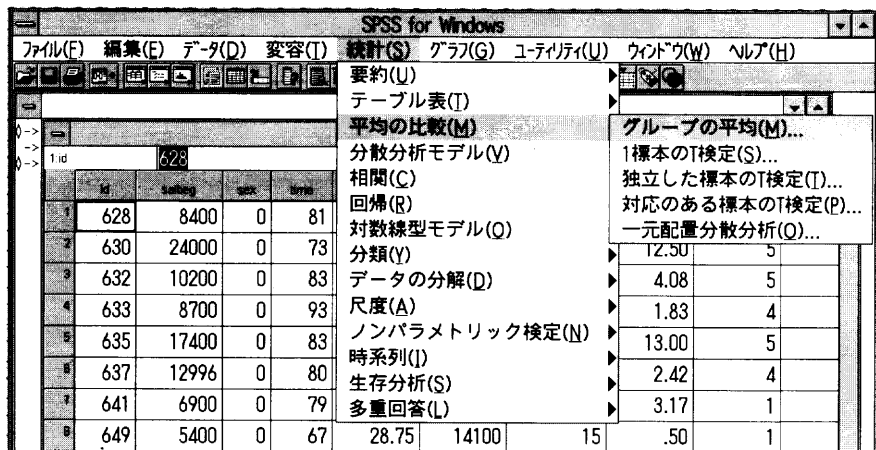


図9. 統計メニューと平均の比較サブメニューのポップアップ.

データの取り扱いに関しては、表計算ソフトとデータベースソフトの持つ便利な機能のほとんどを兼ね備え、しかも操作性が良いという特徴を持つ (図9)。

統計解析作業も直接コマンドを記述し実行するか、もしくはアイコンとダイアログから目的の統計解析手法 (サブプログラム) を選択しコマンドをシンタックスウインドウに貼り付けてからそのコマンドを実行するか、アイコンとダイアログボックスで目的の手法を直接実行できる (図10-1~図10-7)。またその結果でグラフ化の可能なものは、オプションにより強力なグラフィックエンジンを利用して、様々な形式の見栄えの良いグラフとして表示、印刷可能である (図11, 図12, 図13)。SPSS 6.1J for WindowsではBaseに加え Professional Statistics, Advanced Statistics, Tables, Trends, Categories, Lisrel 他のオプションが用意されており、ユーザーの目的に応じて導入できるが各々のオプションに含まれる統計手法は以下の通りである。

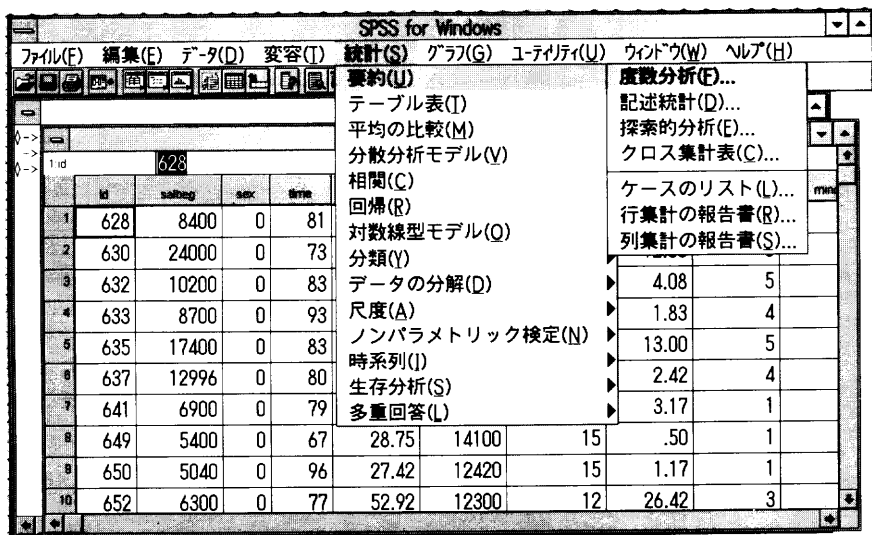


図 10-1. 統計メニューと度数分析サブメニューのポップアップ。

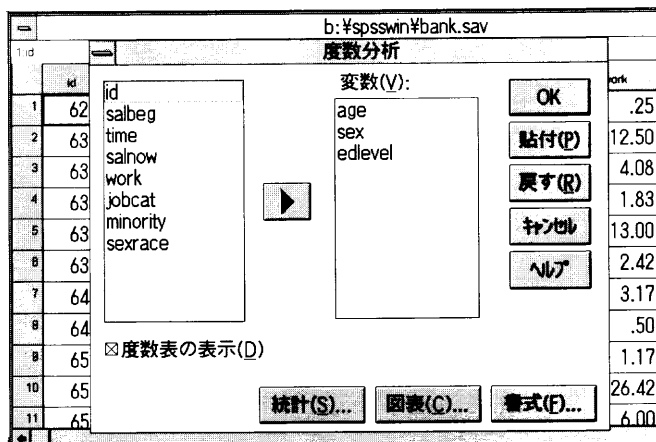


図 10-2. 度数分析のダイアログボックス。

## —SPSS のオプション (SPSS Inc. (1995)) —覧—

## ○SPSS Base System

SPSS Base System は Base Management System, Base Statistics System, Base Graphic System 等で構成される。

## ◇Base (Management) System

Base System は GUI (Graphical User Interface) を重視した SPSS のコントロールセンターで、ファイルやデータの管理とその統計処理、出力・表示等プログラム全体のマネージメントを行う。

## ◇Base Statistics System

Base Statistics は、

Figure 10-3 is a dialog box titled "度数分析: 統計". It is divided into several sections:

- パーセンタイル値**: Includes checkboxes for "四分位(Q)", "同一ケースの(U)", and "パーセンタイル(P)". There is a "グループに分割" checkbox and a list box. Buttons for "追加(A)", "変更(C)", and "削除(R)" are present.
- 中心傾向**: Includes checkboxes for "平均値(M)", "中央値(D)", "最頻値(O)", and "合計値(S)".
- ばらつき**: Includes checkboxes for "標準偏差(T)", "分散(V)", "範囲(N)", "最小値(I)", "最大値(X)", and "標準誤差(E)".
- 分布**: Includes checkboxes for "歪度(W)" and "尖度(K)".
- その他**: Includes a checkbox "値はグループの中央値(L)".

Buttons for "続行", "キャンセル", and "ヘルプ" are located on the right side.

図 10-3. 度数分析: 統計サブダイアログボックス.

Figure 10-4 is a dialog box titled "度数分析: 図表の設定". It is overlaid on a window titled "度数分析" which shows a list of variables: id, salbeg, time, salhow, work, jobcat, minority, sexrace. The dialog box contains:

- 図表の種類**: Radio buttons for "なし(N)", "棒グラフ(B)", and "ヒストグラム(H)". A checkbox for "正規曲線を伴う(W)".
- 軸ラベルの表示**: Radio buttons for "度数(F)" and "パーセンテージ(P)".

Buttons for "続行", "キャンセル", and "ヘルプ" are on the right. Buttons for "統計(S)...", "図表(O)...", and "書式(F)..." are at the bottom.

図 10-4. 度数分析: 図表の設定サブダイアログボックス.

Figure 10-5 is a command window titled "シNTAX1" showing the following syntax:

```

FREQUENCIES
  VARIABLES=age sex edlevel
  /NTILES= 4
  /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN
  /HISTOGRAM .
  
```

図 10-5. 「貼り付け」コマンドボタンによるシNTAXの記述.

出力					
EOLEVEL 就学年数					
Value Label	Value	Frequency	Percent	Valid Percent	Cum. Percent
	8	53	11.2	11.2	11.2
	12	190	40.1	40.1	51.3
	14	6	1.3	1.3	52.5
	15	116	24.5	24.5	77.0
	15	59	12.4	12.4	89.5
	17	11	2.3	2.3	91.8
	19	9	1.9	1.9	93.7
	19	27	5.7	5.7	99.4
	23	2	.4	.4	99.8
	21	1	.2	.2	100.0
	Total	474	100.0	100.0	

Hi-Res Chart # 4: 就学年数のヒストグラム					
Mean	13.492	Median	12.000	Std. dev.	2.885
Minimum	8.000	Maximum	21.000		
Percentile	Value	Percentile	Value	Percentile	Value
25.00	12.000	50.00	12.000	75.00	15.000
Valid cases	474	Missing cases	0		

図 10-6. 度数分析実行結果の出力ウィンドウ。

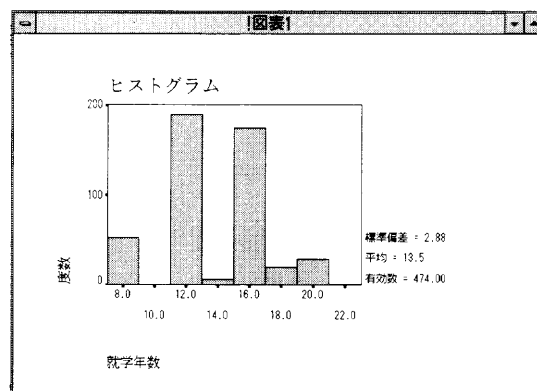


図 10-7. 度数分析実行結果のグラフウィンドウ。

度数表作成 (度数表作成とパーセント値の算出等)

記述統計 (基本統計量の算出等)

データの探索的分析 (データの特性的探索: 正規性の検定やヒストグラム作成等)

クロス表作成 (7 次元以内のクロス表作成と独立性の検定 ( $\chi^2$  検定等))

平均の差の記述 (サブグループ別集計や平均の差の検定 (t 検定), 1 元配置の分散分析 (信頼区間, 等分散の検定, 分散分析, 事後の多重比較等))

多重回答処理 (多重回答データの度数表とクロス表作成等)

分散分析

相関分析 (単純相関, 偏相関)

線形回帰分析 (回帰, 重回帰, 曲線推定等)

ノンパラメトリック検定 (Mann-Whitney,  $\chi^2$ , 二項, ラン, Kolmogorov-Smirnov 他)

リスト作成 (ケースのリスト)

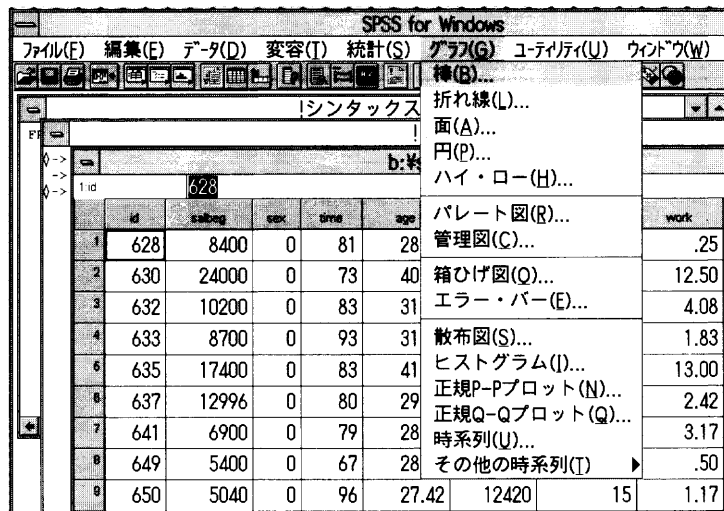


図 11. グラフメニューのポップアップ。

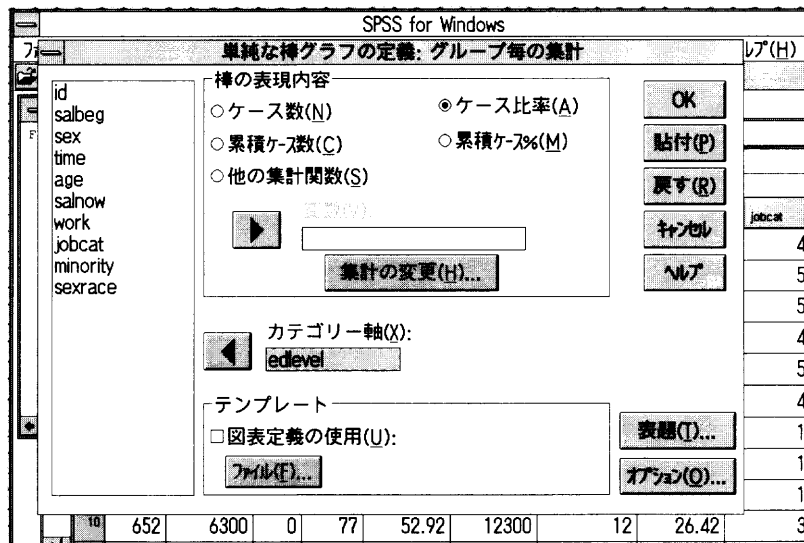


図 12. 棒グラフ定義ダイアログボックス。

レポート作成 (報告書作成)

Graph Engine (System) グラフエンジン (種々のグラフ作成)

で構成されている。

グラフエンジンは、サブプログラムを用いて出力された統計の結果をグラフ化したり、データを直接処理して希望する形のグラフとして表示したり、できあがったグラフを編集して適切な形に変えることが可能である。

通常の一般的な統計処理とその報告書作成のコンピュータツールとしては、このベースシステムのみでも十分対応可能である。

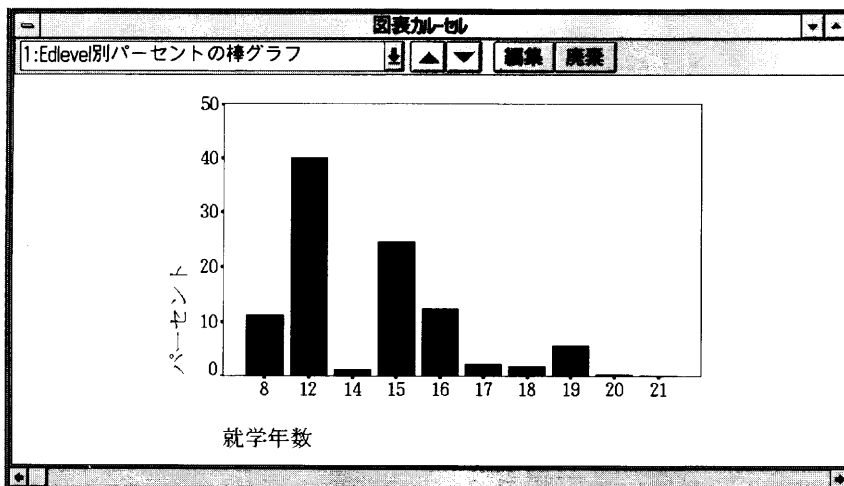


図 13. グラフ出力画面.

## ○SPSS Professional Statistics

Professional Statistics には判別分析, 因子分析, クラスター分析, 大きなファイルのクラスター分析, 距離と近接性の測度, 信頼性分析, 多次元尺度構成法, 最小二乗回帰他, 主としてデータの構造を解析するサブプログラムが含まれている。

## ○SPSS Advanced Statistics

Advanced Statistics は Logistic 線形回帰モデル, 一般化因子の分散分析 (ANOVA), 多変量分散分析 (MANOVA), 非線形回帰モデル, プロビット分析, 生命表分析, Kaplan-Meier 生存曲線分析, Cox の回帰分析など実験計画や生物統計の手法を包括している。

## ○SPSS Tables

Tables は Base システムのクロステーブルや多重回答のクロス表の処理をより高度になすもので, 表の形式にまとめられるデータを様々に修飾し, そのまま印刷原稿に使用可能な品質と内容でテーブル (表) 化するサブプログラムである。

## ○SPSS Trends

Trends は高度な時系列解析の環境を対話形式で与えるサブプログラムで, データのプロット, 平滑化, 成分分解, スペクトル分析, 回帰分析, ARIMA (Auto Regressive Integrated Moving Average) モデル化が行える。季節性因子の推定に用いる。

## ○SPSS Categories

Categories は社会科学で用いられる統計モデルの対象になりにくいカテゴリーデータを取り扱う統計サブプログラムで, 効果的な製品設計を行うための市場調査ツールであるコンジョイント (Conjoint) 分析とカテゴリーデータによって作成された分割表データを符号分析, 非線形主成分分析, 非線形正準相関分析の手法を用い, 行カテゴリーと列カテゴリー間の最大の相関の解析を試みるものである。カテゴリー化された複数の要因の組み合わせのうち, どの組み合わせがもっとも効果的かを推定することができる。

## ○SPSS Lisrel

Lisrel はいくつかの統計モデルで成り立っており, 因子分析をより統計学的に処理しようと

する手法である検証的因子分析（最尤因子分析または測定モデルとも呼ばれる）と、多重指標多変量回帰モデル、構造方程式モデル（逐次モデル、非逐次モデル、多変量回帰、パス解析からなる）、2階因子分析モデル、シンプレックスモデル、多重指標線形構造式モデルが含まれている。共分散構造分析とも呼ばれる。

#### ○SPSS CHAID

CHAIDは $\chi^2$ 検定を探索的に実行するサブプログラムで、3変数以上のカテゴリーデータで構成されたファイルを対象として従属変数に影響を与える独立変数を明らかにする手法である。ファイルをカテゴリーカルなグルーピング変数によって分割した時、従属変数の分布が独立変数で分割されるグループ間で異なるが、これを $\chi^2$ 検定で独立変数で分割されるグループ間の差の強さ、すなわち独立性の高さで評価する。これを各独立変数ごとに実施し、従属変数に対する独立性の強固さの順に並べ替え、いくつかの有意な独立変数で分割される階層ごとに連続的に $\chi^2$ 検定を行い、従属変数と複数の独立変数の関連性を階層的に解析する。従属変数のうける独立変数の影響を構造的に明らかにする独立変数の選択や優先度はユーザーが指定できるモードもある。

#### ○その他

SPSSの新バージョン（Ver. 7.5）ではFamily ProductsとしてExact Test, Neural Connection, allCLEAR, QI analyst, Mapinfo, Teleformが用意（一部英語版のみ）されることになっている。

Exact Testは通常の統計手法が適用不可能な少数例や二項近似で解析不能なデータを厳密に解析するツールである。

Neural Connectionは予想、分類、時系列解析、データ分割をneural analysisの手法を用いて実施するツールであり、アイコンで簡単に解析を進めることができる。

allCLEAR, QI analystは製造業他における品質管理や精度管理のツールとして用いられる。allCLEARはフローチャート作成を容易にし、QI analystは統計的管理が簡単に行える。

MapinfoやTeleformは調査のためのツールでファイルに登録された電話番号に自動的にダイヤルし、聞き取ったデータを即時入力し解析用のデータを作成するツールである。Mapinfoは調査結果を地理的メディア（地図等）を介して管理するGISである。

### 5. SPSSの特徴

どの統計解析ソフトも利点と欠点を併せ持つが、カタログやマニュアルでは予想がつかず、こんな事も出来ないのかと失望する反面、便利な機能を見つけて感心したりもする。

SPSSでは利用者の立場からみて重大な欠点はないと思われるが、余りにもユーザーインターフェイスが良いため、統計に不慣れなあるいは理解していないユーザーがSPSSを使用したとき、間違ったモデルの適用であっても数字で結果が出てくることから、初心者が統計の誤用例を引き起こす可能性を否定できない。現在、手法の適用はSPSSユーザーにゆだねられているが、初心者が誤用しないよう、手法の適用に関するデータチェックステージか、分かりやすいチュートリアルヘルプが望まれる。また、統計があらゆる分野で使用されることから、一般的とはいえないまでもたとえば医学・生物学分野で多用される手法のサブプログラム開発と市場への提供が遅い、もしくは実現されないことは、非常に残念に感ずる。また、企業防衛のためとはいえ、ハードキープロテクションが要求されていることは携帯型PCが普及した現在、良心的ユーザーにとって不便で邪魔な存在でしかない。マニュアルの充実も課題であり、現在オフィシャルに提供されている文法書と実例集の中間的なもの、すなわち文法の記述やアイコン



の選択方法を中心にした実例を用いた操作解説書で、出来るだけシンプルなもの望まれる。現在では、SPSS を利用した統計解説書（新村（1994, 1995）、Norusis（1995）、石村（1995））が数点出回っておりこれで代用も可能であるが、便利な利用法は一般ユーザーでは網羅しきれないので、開発元が適切なマニュアルを提供すべきであろう。

SPSS の利点は一口でいえばオーソドックス、その操作性とコスト・ベネフィットの良さであろう。Windows OS 時代に突入し大幅に GUI を採用したため、統計の基本知識さえ有ればアイコンを選択しポップアップするメニューから目的の操作や解析対象となる変数を選択したり、ダイアログボックスに若干のパラメータを入力するだけで、様々なフォーマットのデータファイルを読み込み、データを整理・加工、解析し、結果を数表やグラフで出力出来る。また、統計の手法によっては基準値や有意水準などユーザーで決定しなくてはならないパラメータも多いが、SPSS で定められているデフォルト値を用いて解析を進めても、デフォルト値が極めて統計学的常識に基づいて設定されているため、多くの場合そのまま利用可能な解析結果を得ることが出来る。正式ユーザーに対する一般的な操作や統計手法に関するサポートも良好で、分からない点をたずねて腹立たしい対応に気分を害することも経験しない。コスト面において、一般ユーザーに対する使用権買い取り市販価格はグラフエンジンと多くの基礎統計を含む SPSS Base System 一セットが十数万円で、個人ユーザーや研究者にとって一時的投資額として高額な感は拭えない。しかしながら、データソースを選ばないこと、数年間にわたり使用することや、特別な人材でなく誰でも操作できることを考えると、ユーザーに対するベネフィット、エフェクトは高い。ソフト業界では同一ソフトの多数導入にたいするサイトライセンス契約の設定が一般的となっており、これが利用できればユーザー側における導入コストも更に削減可能と思われる。また、SPSS Inc. が考える教育的配慮から学生ユーザーに便宜を図るため、利用希望者が若干の制限事項をクリアし学生であることの証明を行えば、手続きにより安価に SPSS を利用できるプログラム（SPSS Inc.（1995））が試験的に用意されている。市販の SPSS と全く同等のソフトウェアが、一般的な大学在学期間に相当する四年間の使用権に限定されていると同時に、マニュアルは付属しないが、SPSS Base System で 2 万円で提供されている。このプログラムは SPSS Professional Statistics, SPSS Advanced Statistics にも別料金で適用されている。

## 6. 結 語

以前 SPSS は Statistical Package for Social Science（社会科学のための統計パッケージ）というフレーズで世に紹介された。それが社会のニーズとソフトウェアの発展により変革し“Statistical Products and Service Solution”すなわち「統計を利用した有効な問題解決」の道具に変身した。現在では“Real Stats, Real Easy”, 「必要な統計手法が、誰でも簡単に！」のキャッチフレーズのもと、開発・提供が行われている。以前は統計とコンピューターと統計ソフトの 3 種の知識・技術が必要だった仕事が、本文に示した通り GUI の高度な利用によってコンピューターと統計の基本的知識さえ有れば誰でもアイコンを選ぶだけで実現できるようになった。情報が氾濫し、情報をうまく利用しなければ生き残れない現在、SPSS のような統計パッケージも今や家電商品となったワードプロセッサと同じレベルで考えなければいけない時代になったのである。

SPSS の Ver. 7.5 ではクロス表作成や因子分析、Logistic Regression 等幾つかのサブプログラムが以前にまして強力な解析が可能となるよう改善され新しくソースが書き換えられた他、toolbar のカスタマイズ、テーブルの形式選択（結果の表の形式が用意されている十数種のパターンから自由に選択し、ワープロに貼り付けて印刷原稿にすることが出来る）、Internet への

対応, Windows 95 の 32 ビットモードに完全対応 (ファイルのロングネーム等) など機能が強化され, よりユーザーに便利なものとなっている。また, データファイルを作成するツールである Data Entry for Win. や共分散構造解析 Lisrel の新バージョン AMOS 等様々な新プロダクトが用意されている。さらに, Ver. 8, 9 に向けデータマネージメント (Client-server 対応等), 統計サブプログラム (新手法の取り込み), グラフィックエンジン (Drug & drop 等) 大幅な開発・改定が進行中である。これが実現すればますます便利な道具として統計ソフトが身近な存在として利用されるようになるだろう。

統計ソフトの発展により, 統計が身近になり, 研究や意志決定のための有用な道具として社会的認知を得る日が間近であることを信じて止まない。

## 謝 辞

本稿をまとめるに当たり, SPSS 社の高柳良太氏, 坂本保友氏, 村田悦子氏, 星名良一氏より貴重な参考資料とご意見をいただいた。又, 日本電気, 日立製作所より資料をご提供いただいた。ここに深謝する。

## 参 考 文 献

- 石村貞夫 (1995). 『SPSS による統計処理の手順』, 東京図書, 東京。
- 三宅一郎 (1973). 『社会科学のための統計パッケージ』, 東洋経済新報社, 東京。
- 三宅一郎, 山本嘉一郎 (1976). 『SPSS 統計パッケージ 基礎編』, 東洋経済新報社, 東京。
- 三宅一郎, 山本嘉一郎 (1986). 『新版 SPSSx I 基礎編』, 東洋経済新報社, 東京。
- 三宅一郎, 中野嘉弘, 水野鉄司, 山本嘉一郎 (1977). 『SPSS 統計パッケージ 解析編』, 東洋経済新報社, 東京。
- 三宅一郎, 山本嘉一郎, 垂水共之, 白倉幸男, 小野寺孝義 (1991). 『新版 SPSSx III 解析編 2』, 東洋経済新報社, 東京。
- Norusis, Maria J. (1994). *SPSS for Windows. Base System User's Guide Release 6.1*, SPSS Inc., Chicago.
- Norusis, Maria J. (1995). 『SPSS による統計学入門』 (山本嘉一郎 他 訳), 東洋経済新報社, 東京。
- 司馬正次 (1977). 『データ解析入門 SPSS への招待』, 東洋経済新報社, 東京。
- 司馬正次 (1989). 『やさしいデータ解析 SPSSx による』, 東洋経済新報社, 東京。
- 新村秀一 (1994). 『SPSS for Windows 入門』, 丸善, 東京。
- 新村秀一 (1995). 『パソコンによるデータ解析』, 講談社, 東京。
- SPSS Inc. (1994). *SPSS 6.1 Syntax Reference Guide*, SPSS Inc., Chicago.
- SPSS Inc. (1995). SPSS 社内部資料, SPSS Inc., 東京。
- 垂水共之, 西脇二一, 石田千代子, 小野寺孝義 (1990). 『新版 SPSSx II 解析編 1』, 東洋経済新報社, 東京。
- 山本嘉一郎, 吉村 英, 竹村和久 (1991). 『パソコン SPSS 基礎編』, 東洋経済新報社, 東京。

## Real Stats, Real Easy—SPSS, a Useful Statistics Products—

Masayuki Hayashi

(Department of Demography and Health Statistics, The Institute of Public Health)

SPSS is one of famous softwares for statistics in the world. User interface and reliability are very good as recognized by many scientists.

With the statistics in SPSS, users always get a right solution for the job.

SPSS easily imports data from a variety of sources. Of course, new data can be entered to SPSS spread sheet, then, users can choose from basic to sophisticated statistics and graphics with in-depth options. SPSS's full-featured statistics help its users to discover immediately if its results are significant or if differences are random.

SPSS has most of statistical models which we use generally. And users can select and make running statistical tools by choosing eyecons (icons) displayed on a menu bar or a tool bar of SPSS. These products are easy to use and easy to learn without special knowledge and technique. In the case of cross tables, there are many kind of table formats according to an object of users. SPSS makes user's data exploration faster, easier and more interesting.

There are many problems for SPSS users in Japan. For example, a tutorial help for beginners are not enough, a hard key is required in Asian area, new statistical method (eg. medical or bio statistics) was difficult to set the package and so on. Nevertheless, SPSS is almost good to prepare data and to analyse quickly. It can analyse and examine easily and present results clearly. And SPSS has good cost-benefit and cost-effect in comparison to similar kind of statistical package of PC.

Now, we are in many 'Information'. We have to utilize information and make our decision with a statistical results of them. For a decision making, SPSS is the first choice for statistics users.