

SPSS統計パッケージ

I 基礎編

三宅一郎
山本嘉一郎

SPSS統計パッケージ

I 基礎編

三宅一郎 著
山本嘉一郎

東洋経済新報社

著者紹介

みやけ いちろう
三宅 一郎

1931年 神戸に生まれる。京都大学法学部卒業，同大学大学院，同大学
人文科学研究所講師を経て，現在同志社大学法学部教授。

専攻 政治学〔政治意識論，政治過程論〕

著書 『異なるレベルの選挙における投票行動の研究』（共著）創文社，
『社会科学のための統計パッケージ』東洋経済新報社。

住所 西宮市高松町15-16

やまもと か いちろう
山本 嘉一郎

1943年 宮崎県に生まれる。京都大学理学部卒業，同大学大学院理学研
究科，博士課程修了（理学博士）。

専攻 地質鉱物学

住所 豊中市緑丘1-24-5

SPSS 統計パッケージⅠ 基礎編

定価 3500 円

昭和51年11月15日 第1刷発行

昭和59年7月5日 第10刷発行

著者 三宅一郎／山本嘉一郎

発行者 高柳 弘

発行所 〒103 東京都中央区日本橋本石町1の4 東洋経済新報社

電話 編集 03(246)5661・販売 03(246)5467 振替 東京3-6518

本書の一部または全部の複写・複製・転載・磁気媒体への入力等を禁じます。これ
らの許諾については，小社（電話03-246-5634）までご照会ください。

© 1976 〈検印省略〉 落丁・乱丁本はお取替えいたします。

Printed in Japan ISBN 4-492-47018-2

は し が き

本書は、前著『社会科学における統計パッケージ』の改訂増補版である。1973年に出版された前著は、アメリカのスタンフォード大学で生まれ、シカゴ大学で育った SPSS (Statistical Package for the Social Sciences) 第4版に、数量化理論などいくつかの国産プログラムを加えた京大第4版に基づいていた (SPSS およびその京大版の歴史については、前著の「序にかえて」で述べた)。その後、SPSS はシカゴにおいても、わが国においてもレベルアップを重ねたが、特に第5版から第6版へのレベルアップにあたり、命令語の変更を含む大幅な改訂がなされたうえ、これを機にわが国では、クラスター分析などの新しいサブプログラムを独自に編入した。このため、前著は SPSS の解説書として不十分なものになったので、最新の SPSS 第6版に基づき、ほとんど新しく書き改めた改訂版を出版することにした。

前著出版当初には、SPSS を利用できる計算機センターは、FACOM 230/60 (当時)をもつ京都大学と北海道大学の兩大型計算機センターのみであったが、やがて同じ機種 of 計算機を備える名古屋、九州兩大学の大型計算機センターが加わり、本年度からは東京大学大型計算機センター用の HITAC 8700/8800 (OS7)版が公開され、さらには HITAC 中型機用のミニ SPSS も、同志社大学と神戸大学経済経営研究所の計算センターでサービスを始めているし、FACOM 中型機用のミニ SPSS のテスト版も完成している。また、ACOS 900, TOSBAC 5600, FACOM M-190 版 SPSS の開発作業が進行中である。これらの作業が完了すると、全国 of 大型計算機センターはもちろん、多くの大学、研究所の計算センターでも SPSS の利用が可能になるはずである。

この間、利用者也年々増加してきたが、このパッケージ開発当初の対象であった社会科学系の研究者ばかりでなく、最近では理科系の利用者がふえているのが目立つ。この事実は、SPSS のすぐれたファイル管理やデータ処理機能が広

く知られるにいたったためでもあろうが、また統計手法そのものの学際化が著しく進んでいるという現状の反映でもあろう。本書の書名を変更して「社会科学における」を除去し、『SPSS 統計パッケージ』としたのは、こういう事情によるものである。

書名だけでなく、本書では記述内容についても多くの改訂を行なった。特に大きな改訂は次の二つである。まず、前著では統計教科書に譲ってほとんど説明をしなかった、統計サブプログラムの計算手法の具体的内容や検定の説明を、きわめて簡単ではあるが、記述するよう努めた。第2に、SPSS コントロールカードの文法エラーや実行時のエラーの解釈を容易にするために、エラー番号を本文中に挿入し、巻末にエラー番号索引を付した。

こうした改訂と、4版以来編入された機能とサブプログラムの解説の追加のために、本書のページ数は増加する一方であり、したがって出版が遅れるおそれがあった。そこで、コントロールカードと基礎統計サブプログラムの解説を「Ⅰ 基礎編」とし、その他の多変量解析サブプログラムを「Ⅱ 解析編」とする2分冊に分けて刊行することにした。この点、読者のご寛容を乞うものである。

SPSS は使いやすさの点でも、統計サブプログラムの内容の点でも、最もすぐれたパッケージの一つであるが、もちろん欠点もまだ残されているし、新しい統計分析手法が次から次へと開発されている現況からみると、統計サブプログラムの種類もこれで十分だとはいえない。今後も、エラーの訂正、新しいサブプログラムの編入に努力するつもりである。シカゴでは AID, Linear Log Model, Least Square などのサブプログラムを準備中というし、わが国でも、多次元尺度構成サブプログラムの編入や、他のすぐれたパッケージとの連結などを予定している。なお、これと平行して、会話型 SPSS の開発が進行している。シカゴではすでにテスト版がつくられているが、われわれも会話型 SPSS に関心をもっており、近くその研究を開始する。

われわれの小さな努力が、全国の研究者の計算機利用を助け、ひいてはわが国の科学研究の発展に少しでも貢献できれば、われわれにとってこれにまさる喜びはない。だが、SPSS のような汎用パッケージの整備や発展は、ひとり企画者やプログラマーの力によるものでは決してない。利用者の協力に多くを負っていることは、SPSS のこれまでの歴史が示すとおりである。利用者の貢献

は、エラーの発見やレベルアップのための示唆をはじめ、新しいサブプログラムの作成編入にまで及んでいる。一般に、利用者のコミュニティの問題関心と統計手法が、パッケージに形を与えそれを育てる。利用者との対話によるパッケージの発展が、推進者としての著者たちの望みである。利用者の積極的な協力をせつに期待している。

SPSS 京大版の開発と前著の出版にあたり、SPSS の原著者をはじめ、多くの人々にご指導、ご援助をいただいたことは、前著に感謝をもって記したとおりである。その後、SPSS のレベルアップと適用機種 of 拡大作業は、京都大学における橋本弘蔵氏と三宅の小さなグループの仕事から、全国の各地にいくつかの拠点をもつ複合グループの共同プロジェクトへと発展していった。これに伴い、SPSS の書き替え作業やプログラムテスト作業に関し、ご援助やご協力をいただいた人々は、かつての数倍にも達した。また直接作業に携わらなくとも、SPSS の一般利用のため鋭意尽力下さった方々も多かった。SPSS の東京大学大型計算機センターへの導入に代表責任者として努力され、本年7月急逝された故池内一氏(東大新聞研教授)もそのお一人である。このような多くの方々にこの紙面で感謝の辞を述べることは、ほとんど不可能である。ここでは、直接作業グループに参加された人々のお名前のみをあげさせていただいて、感謝の表明としたい(敬称は省略させていただく)。

FACOM 230-60/75 第6版へのレベルアップ作業

京都大学 飯田記子, 橋本弘蔵, 村尾義和, 山本嘉一郎, 西脇二一, 岡本安晴

北海道大学 石川恒, 司馬正次, 須藤研, 山口淳一, 土井時久, 長南史雄,
吉川経教

北海道教育大学 中野嘉弘, 岡本次郎, 高木富士夫

北海道工業大学 伊藤正義, 島 貢

札幌医科大学 竹川忠男

名古屋大学 水野欽司, 西本史雄, 佐々木雅子

HITAC 8700/8800 第5版の作成および第6版へのレベルアップ

東京大学 宮下允, 石田千代子

広島大学 垂水共之

HITAC 8350 ミニ SPSS の開発

神戸大学 定道 宏

同志社大学 安田 聖

FACOM 230-38/45S ミニ SPSS の開発

関西学院大学 雄山真弓

明治大学 大竹政光

法政大学 蜂谷 博

東京都立大学 西村弘之

この作業は、京都、北海道、名古屋各大学大型計算機センター、神戸大学経済経営研究所機械計算室、関西学院大学、明治大学各計算センターの開発計画の一部であり、各センターから多大の便宜を得た。これらの援助がなければ、今日の SPSS は存在しえなかった。各センターの教官、教員、職員、担当システムエンジニア諸氏のご支援、ご指導に厚く謝意を表したい。また、昭和 50、51 両年度文部省科学研究費補助金(総合研究 A「SPSS のレベルアップと適用機種拡大」)によって、財政的に支持されたことも特記しておかねばならない。

最後になったが、国産機種版の SPSS および本書は SPSS の原著者たち (Professor Norman H. Nie, Mr. C. Hadlai Hull, Ms. Jean G. Jenkins, Ms. Karin Steinbrenner, Professor Dale H. Bent) に負うところが大きいのはもちろんのことであるが、わが国でのサブプログラムの編入にあたっては、BMD の開発者、Rome Air Development Center, Imperial College Computer Center, Dr. J. E. Klovian, Dr. J. M. Parks, および駒沢勉氏らに多くを負っている。ここで改めて深い感謝をささげさせていただく。これらすべての方々のご好意があってはじめて、国産機種版の SPSS はありえたわけであるが、その書き直し作業や編入作業上での誤り、あるいは本書における解説に誤りがあるとすれば、その責任はすべて著者らに帰せられるべきはいうまでもない。

この本の出版にあたって、今回も東洋経済新報社の丸山常喜氏に終始お世話になった。お礼を申しあげる。

昭和 51 年 9 月

著者代表 三宅一郎

本書の使い方

この本の使い方について、SPSS の利用経験者と未経験者によって違った注意が必要と思われるし、また、利用する計算機センターの機種によっても、読み方に差が生じるので、ごく簡単ながらはじめに注記しておきたい。

1. SPSS 利用経験者と未経験者

1.1 利用未経験者への注意

(1) 特定のサブプログラムのみ利用しようとする人も、一度は必ず第1章より読み始めてほしい。

(2) 文中に多くの割注がはいっているが、SPSS の機能を概観するために、この本に目を通すときは、この割注を飛ばして読んでもよい。しかしサブプログラムの利用のためには、割注に書かれた注意は非常に重要である。

(3) ページ数節減のため、一度説明したことの繰返しはできるだけ避けて、問題を叙述した節あるいは項の参照を⇒印で指示したが、必ずこの指示に従ってほしい。

(4) コントロールカードの文法エラーやデータエラーが発見されるとエラーメッセージが印刷されるが、このメッセージの意味を理解する手がかりとして、エラー番号索引を巻末に付したので利用をすすめる。また、京都大学大型計算機センターより、「SPSSエラーメッセージ集」が刊行されているので、これも合わせて参照するのが望ましい。

1.2 経験者のために

(1) SPSS 利用経験者、特に前著『社会科学のための統計パッケージ』をマニュアルとして利用してこられた人のために、補遺Ⅰに「第4版から第6版へ」を収録したので、まずこれに目を通して、前著と本書の間の違いについてのおおよそのところをつかんでほしい。

(2) 補遺Ⅵ「コントロールカードの一般書式、オプション、統計一覧表」も

経験者の便宜のために編集したものである。

(3) エラーメッセージの解説にあたり、上記 1.1 (4) の「SPSS エラーメッセージ集」も併用されたい。

(4) SPSS の入力データの準備や出力の再加工のために利用できる、いくつかのプログラムを補遺Ⅲに紹介しておいたので利用されたい。

2. 利用機種による SPSS 版種の違い

現在、わが国で利用できる SPSS および目下開発中の SPSS の版種は、大型計算機用 (大型 SPSS または SPSSH) か中型計算機用 (ミニ SPSS または SPSSG) が、ワード計算機用かバイト計算機用かという二つの基準によって、次の表のように 3 種に大別できる。

	大型 SPSS (SPSSH)	ミニ SPSS (SPSSG)
ワード計算機用	FACOM 230-60/75 ACOS 900 TOSBAC 5600	
バイト計算機用	IBM 370 HITAC 8700/8800 FACOM M-190	HITAC 8350 FACOM 230- 35/38/45S

(注) IBM 以外の外国機種版は省略した。

本書は大型ワード計算機用 (特に FACOM 230-60/75) の SPSS 第 6 版を中心に説明しているが、他の版種にも注意を払い、版種により重大な相異があるときは、脚注をつけて記述した。他の版種の利用者は脚注を見逃さないように注意してほしい。

ワード計算機用版とバイト計算機用版とのおもな違いは、必要作業領域の計算式にある。本書で記述する作業領域 (WORKSPACE) はワード単位であるから、バイト計算機用にはこれを 4 倍するとよい。

ミニ SPSS 利用者のためには、特に補遺Ⅱに大型 SPSS とミニ SPSS の違いをまとめたので、まずこれに目を通されたい。

しかしながら、同じ版種の SPSS といっても、計算機センターごとにその内容が異なる場合も少なくない。特にその利用に必要なジョブ制御文は、まったく同一機種の計算機をもつ計算機センター間でも異なるのが普通であるから、初めての利用者は、その SPSS の内容およびジョブ制御文について、あらかじめ計算機センターに問い合わせる必要があろう。

「Ⅱ 解析編」の主要目次

- 22 PARTIAL CORR : 偏相関係数
- 23 REGRESSION : 重回帰分析
- 24 ANOVA : 多元配置分散・共分散分析
- 25 ONEWAY : 一元配置分散分析
- 26 DISCRIMINANT : 判別分析
- 27 CANCORR : 正準相関分析
- 28 FACTOR : 因子分析
- 29 QFACTOR : Qモード因子分析
- 30 HAYASI 1 : 数量化理論 1 類
- 31 HAYASI 2 : 数量化理論 2 類
- 32 HAYASI 3 : 数量化理論 3 類
- 33 HAYASI 4 : 数量化理論 4 類
- 34 CLUSTER : クラスター分析
- 35 NONLINEAR MAP : 非線形マッピング
- 36 RELIABILITY : 信頼性係数

目 次

は し が き

本書の使い方

I SPSS の特性とデータ

1	SPSS システムの特性	3
1. 1	プログラムの有機的連結	3
1. 2	SPSS 言語と自由書式制のコントロールカード	3
1. 3	データ情報を内蔵するデータファイルと情報検索	4
1. 4	サブファイル構造	5
1. 5	データ変容の機能	5
1. 6	欠損値の処理	6
1. 7	ノンメトリックな統計機能の充実	6
1. 8	外部媒体へのデータ出力	8
1. 9	大量データの処理能力	8
1.10	プログラム制限の柔軟性	9
2	データの構造とコーディング	11
2. 1	ケースと変数	11
2. 2	データのコーディング	12
2. 2. 1	変数の型と変数値	12
2. 2. 2	入力データの編成	13
2. 2. 3	ケース番号とカード番号	14

2. 2. 4 データのクリーニング.....	15
2. 3 サブファイルの機能と構造	16
2. 3. 1 サブファイルの機能.....	16
2. 3. 2 サブファイルの構造.....	17
2. 4 カード以外の媒体の利用	18
2. 5 入力データ量の制限.....	18

II SPSS のコントロールカード

3 SPSS のコントロールカード	23
3. 1 コントロールカードの一般書式と基本的語彙	23
3. 1. 1 命 令 領 域.....	23
3. 1. 2 明細指示領域.....	26
3. 1. 3 コントロールカードに関する一般的規則.....	28
3. 2 コントロールカードの種類	29
3. 3 コントロールカードとエラーメッセージの印刷	30
3. 4 コントロールカード書式の説明のための凡例	31
4 SPSS ファイルの定義：データ定義カード.....	32
4. 1 FILE NAME カード.....	33
4. 2 VARIABLE LIST カード.....	34
4. 3 SUBFILE LIST カード	36
4. 4 N OF CASES カード	37
4. 5 INPUT MEDIUM カード	38
4. 6 INPUT FORMAT カード.....	38
4. 6. 1 データケースの編成：固定コラム，自由領域， バイナリー書式.....	38
4. 6. 2 固定コラム書式における書式文.....	39
4. 6. 3 バイナリー書式における書式文.....	41
4. 7 DATA LIST カード：VARIABLE LIST と INPUT FORMAT カードの代替機能.....	42

4. 7. 1	固定コラム書式のための DATA LIST カード	42
4. 7. 2	パイナリー書式のための DATA LIST カード	44
4. 8	MISSING VALUES カード	45
4. 8. 1	変数リスト	46
4. 8. 2	キーワード ALL	47
4. 9	VALUE LABELS カード	47
4.10	PRINT FORMATS カード	49
4.11	VAR LABELS カード	50
4.12	DOCUMENT カード	51
4.13	データ定義カードの順位と例	52
5	課題定義	53
5. 1	TASK NAME カード	53
5. 2	手続きカード	53
5. 3	OPTIONS カード	54
5. 4	STATISTICS カード	55
5. 5	RUN SUBFILES カード	56
5. 6	RAW OUTPUT UNIT カード	57
5. 7	課題定義カードの総合例	58
6	run コントロール	60
6. 1	RUN NAME カード	60
6. 2	EDIT カード	60
6. 3	READ INPUT DATA カード	62
6. 4	FINISH カード	62
6. 5	PRINT BACK カード	62
6. 6	NUMBERED カード	63
6. 7	PAGESIZE カード	63
6. 8	COMMENT カード	64

7	SPSS ファイルの作成, 処理, 保存, 検索	65
7. 1	SPSS ファイルの作成と処理: 初期入力データの処理	65
7. 2	SAVE FILE カード: SPSS システムファイルの 作成と保存	67
7. 3	GET FILE カード: SPSS システムファイルから 入力したデータの処理	69
7. 4	SPSS ファイルよりの情報検索	70
7. 4. 1	LIST FILEINFO カード: データ定義情報の検索(1)	70
7. 4. 2	WRITE FILEINFO カード: データ定義情報の検索(2)	71
7. 4. 3	LIST CASES カード: データ出力(1)	72
7. 4. 4	WRITE CASES カード: データ出力(2)	75
8	データ変容カード: 再コードと変数変換	78
8. 1	RECODE カード: 変数の再コード	79
8. 1. 1	*RECODE カード: 一時的再コード	81
8. 1. 2	RECODE および *RECODE カードの制限	81
8. 2	COUNT カード: インデックス作成	82
8. 2. 1	*COUNT カード: 一時的インデックス作成	83
8. 2. 2	COUNT カードおよび *COUNT カードの制限	83
8. 3	COMPUTE カードと IF カード: 変数変換	83
8. 3. 1	COMPUTE カード: 算術式による変数の計算	84
8. 3. 2	IF カード: 条件文による変数変換	85
8. 3. 3	IF カードの省記法	87
8. 3. 4	*COMPUTE カードと *IF カード: 一時的変数変換	87
8. 3. 5	COMPUTE カードと IF カードの制限	87
8. 4	DO REPEAT と END REPEAT カード: 変数変換の繰返し記法	88
8. 4. 1	DO REPEAT, END REPEAT カードの制限	89
8. 5	ALLOCATE カード: データ変容のための作業領域 (コア)の確保	90

8. 6	ASSIGN MISSING カード：データ変容に伴う欠損 値の処置と初期値の設定	91
8. 6. 1	欠損値をもつ変数を使ったデータ変容	91
8. 6. 2	新設変数の初期値設定	91
8. 7	新設または変換変数のためのその他のデータ定義	92
8. 8	データ変容のためのカード順位と構成例	92
9	データ選別カード：ケースの選別とウェイトづけ	94
9. 1	SAMPLE カード：ランダムサンプルの創出	94
9. 2	SELECT IF カード：ケースの選別	95
9. 3	WEIGHT カード：ケースのウェイトづけ	96
10	ファイル変容カード：システムファイルの変容	98
10. 1	DELETE VARS と KEEP VARS カード：変数の 削除と保存	99
10. 2	ADD VARIABLES と ADD DATA LIST カード： システムファイルへの変数の追加	100
10. 2. 1	ADD VARIABLES カードによる変数の追加	101
10. 2. 2	ADD DATA LIST カードによる変数の追加	104
10. 3	REORDER VARS カード：システムファイル内の 変数順序の変更	104
10. 4	MERGE FILES カード：システムファイルの併合	105
10. 5	SELECT IF, DELETE SUBFILES カード：ケー スまたはサブファイルの削除	106
10. 6	ADD CASES, ADD SUBFILES カード：ケース またはサブファイルの追加	107
10. 6. 1	ADD CASES カード	107
10. 6. 2	ADD SUBFILES カード	108
10. 7	SORT CASES カード：ケース順序の変更	109

10. 8	サブファイル構造の変更	111
10. 9	UPDATE IF カード：ケースの取替え	111
10.10	データ定義情報の改訂	113
11	アーカイブの作成，処理，検索	115
11. 1	GET ARCHIVE カード：大量変数ファイルの入力	116
11. 2	SAVE ARCHIVE カード：アーカイブの作成	117
11. 3	LIST ARCHINFO カード：アーカイブファイル 情報の検索	118
11. 4	アーカイブ機能の制限	119

Ⅲ 基礎統計のためのサブプログラム

12	記述統計と度数分布	123
12. 1	記述統計の計算法	123
12. 2	CONDESCRIPTIVE：記述統計	125
12. 2. 1	CONDESCRIPTIVE の手続きカード	125
12. 2. 2	CONDESCRIPTIVE のオプション	125
12. 2. 3	CONDESCRIPTIVE の統計	127
12. 2. 4	CONDESCRIPTIVE のプログラム上の制限	127
12. 2. 5	CONDESCRIPTIVE カードと結果出力例	127
12. 3	FREQUENCIES：度数分布とヒストグラムの作成	130
12. 3. 1	FREQUENCIES の手続きカード	130
12. 3. 1. 1	一般モードの書式	130
12. 3. 1. 2	整数モードの書式	131
12. 3. 2	FREQUENCIES のオプション	132
12. 3. 3	FREQUENCIES の統計	133
12. 3. 4	FREQUENCIES のプログラム上の制限	133
12. 3. 5	FREQUENCIES カードと結果出力例	134

13	AGGREGATE : アグリゲート・データの産出	138
13. 1	AGGREGATE の手続きカード	139
13. 2	出力ファイルの内容と形式	141
13. 3	AGGREGATE のオプション	141
13. 4	AGGREGATE の追加統計	142
13. 5	AGGREGATE の制限	142
13. 6	AGGREGATE カードの結果出力例	143
14	サブグループ別記述統計	145
14. 1	BREAKDOWN : サブグループ記述	145
14. 1. 1	統計量の計算法	146
14. 1. 2	BREAKDOWN の手続きカード	148
14. 1. 2. 1	一般モードの書式	148
14. 1. 2. 2	整数モードの書式	148
14. 1. 3	BREAKDOWN のオプション	149
14. 1. 4	BREAKDOWN の追加統計	149
14. 1. 5	BREAKDOWN のプログラム上の制限	150
14. 1. 6	CROSSBREAK : クロス表仕様の BREAKDOWN	151
14. 1. 7	CROSSBREAK 用のオプション	151
14. 1. 8	CROSSBREAK 用の追加統計	151
14. 1. 9	CROSSBREAK の制限	152
14. 1. 10	BREAKDOWN カードと結果出力例	152
14. 2	DATA PATTERN : データパターン記述	156
14. 2. 1	DATA PATTERN の手続きカード	156
14. 2. 2	DATA PATTERN のオプション	158
14. 2. 3	DATA PATTERN の追加統計	159
14. 2. 4	DATA PATTERN のプログラム上の制限	159
14. 2. 5	DATA PATTERN カードと結果出力例	160
15	T-TEST : 二つの平均の差の検定	164