

Title	統計パッケージSPSSからSPSSXへの移行について
Author(s)	中島, 重雄
Citation	大阪大学大型計算機センターニュース. 1990, 79, p. 61-68
Version Type	VoR
URL	https://hdl.handle.net/11094/65901
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

統計パッケージSPSSからSPSSXへの移行について

システム管理掛 中島 重雄

現在、当センターではSPSSとSPSSX両方のサービスを行っていますが今年度をもって、SPSSを廃止いたします。これに伴いSPSS（以下旧版と呼ぶ）を利用されていた方はSPSSX（以下X版と呼ぶ）に移行をおこなっていただく必要があります。ここで、移行の際の要点について述べます。

1. JCL（ジョブ制御言語）の変更

\$CPROC文をSPSSからSPSSXに変更する必要があります。

【例】\$CPROC文

```
<旧版>
  1      8      16カラム
  $      CPROC  SPSS,,5,230K,-4K,5K
<X版>
  $      CPROC  SPSSX,,5,230K,-4K,5K
```

2. コマンドの書式

旧版ではサブコマンドや明細指定は16カラム以後に記述する必要がありましたが、X版は1カラム以外ならどこでも記述できます。ただし、主コマンドは必ず1カラムから始めなくてはなりません。

【例】コマンド書式

```
<旧版>
  1              16カラム
  FREQUENCIES   GENERAL=ALL
<X版>
  FREQUENCIES VARIABLES=ALL
```

3. データの読み込み

旧版とX版との最も重要な変更としてはデータの読み込みだと思います。変更点は次に挙げるものです。

- ① 旧版で使用していた VARIABLE LIST, INPUT FORMAT, INPUT MEDIUMが廃止され、これらに代わりDATA LISTでおこないます。
- ② READ INPUT DATAが廃止されBEGIN DATAでおこないます。

これに伴い、変更が必要となりますが次のようなコマンドの組合せにより読み込みをおこないます。

1) DATA LISTとFILE HANDLEコマンドによる読み込み

特徴としては、データ読み込みで扱えるレコードサイズが大きい（最大4096バイト/レコード）ことや、データの書き込み、またシステムファイルの読み込み、書き込みがおこなえることです。

2) DATA LISTと BEGIN DATAコマンドによる読み込み

特徴としては使用方法が簡単であるのですが、データの読み込みしか出来ず、（システムファイルの読み込みは不可）また、レコードサイズが80バイト以内であることです。

つぎに、この2種類のデータ読み込みについて述べます。

3-1. DATA LISTとFILE HANDLEコマンドによる読み込み

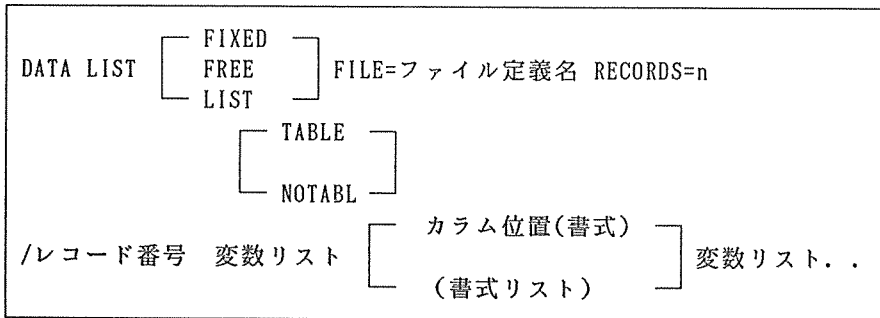
① DATA LISTコマンドでは、入力データファイル、入力変数名とそのファイルでの記録位置などの指定をおこないます。

② FILE HANDLEコマンドでは、DATA LISTコマンドで指定したファイル定義名の指定をおこなったりファイルの書式等を指定します。

また、FILE HANDLEコマンドでは、“外部ファイル名指定”という指定がありますがこれは、\$PRMFL文(JCL)が不用なため便利ではあるのですが反面、ファイルのビジーチェック等がジョブ投入時の段階ではおこなわれません。この理由によりここでは“外部ファイル名指定”については詳しく述べません。少しめんどろですが\$PRMFL文を使った説明をおこないます。

※ FILE HANDLEコマンドについての詳しいことは\$TEBIKIJコマンドで“SPSSX利用の手引”や速報NO. 164(88年11月24日)をご覧ください。

【DATA LISTの一般書式】



《キーワード》

FIXED 固定書式(規定値). 各変数は同一レコード, 同一位置に記録

FREE 自由書式. 変数の記録位置を限定しない. ただし, 変数間はブランクかコンマで区切る

LIST 自由書式で1ケース1レコード

《FILE=ファイル定義名》

ファイル定義名はFILE HANDLE文で指定する必要があります. BEGIN DATAで読み込みをおこなう場合は不必要.

《TABLE, NOTABLEサブコマンド》

変数定義の一覧表出力の指示(規定値TABLE)

《RECORDS=n》

nには1ケースのレコード数を指定. 規定値は1.

《レコード番号 変数リスト. . .》

変数の定義をおこなうレコード番号等の指定.(詳細と例を後述)

【FILE HANDLEコマンド】

FILE HANDLEコマンドにはいくつかの指定方法がありますがここではその内の1つについてのみ説明します。

FILE HANDLE ファイル定義名/DEVICE=装置名/RECFORM=書式/LRECL=レコード長

ファイル定義名 . . . 英文字で8文字以内
装置番号 . . . 装置番号を指定（\$PRMFL文の装置番号と一致）
書式 . . . F, U, SYSのいずれか
F は書式付き
U は書式なし
SYS はシステムファイルの場合
（省略時はF）
レコード長 . . . レコード長（バイト単位）を指定する
最大は4096
規定値は書式付きの場合80バイト，書式無しの場合
512バイト）

【\$PRMFL文】

1 8 16桁
\$ PRMFL 装置番号，割当タイプ，アクセスモード，修飾ファイル名

装置番号 . . . ファイルコードとも呼ばれ2桁の数字を指定
割当タイプ . . . R：読み込み
W：書き込み
アクセスモード . . . S：シーケンシャル
R：ランダム
修飾ファイル名 . . . 利用者番号/ファイル名

3-1-1 固定書式と自由書式

データの読み込みで，ファイル内の変数位置が同じであると固定書式（固定カラム書式）といい，これに対して変数の順序さえ守られていれば同じ変数位置でなくてよいものを自由書式（自由領域書式）といいます。

1) 固定書式

つぎに，固定書式の場合のDATA LISTの説明をおこないます。ここではデータの例として1ケースが1レコードで，番号（3桁），性別（1桁），年齢（2桁），変数T1～T5が正数部2桁，小数部1桁で変数間のブランクはないものとします。

【例1】 DATA LISTコマンド

```
<旧版>
VARIABLE LIST  NUM SEX AGE T1 TO T5
INPUT FORMAT   FIXED(F3.0,F1.0,F2.0,5F4.1)
<X版>
① DATA LIST FILE=INFILE1 /NUM 1-3 SEX 4 AGE 5-6 T1 TO T5 7-26(1)
   または、
② DATA LIST FILE=INFILE1 /NUM SEX AGE T1 TO T5(F3,F1,F2,5F4.1)
```

ここで、T1 TO T5 は
T1 7-10(1) T2 11-14(1) T3 15-18(1) T4 19-22(1) T5 23-26(1)
と同じであり(1)は1桁の小数点を示します。文字型のデータであれば(A)
で示します。例 DATA LIST FILE=INFILE1 / NAME 1-20(A) AGE 21-22
①はデータのカラム位置を指定。NUMは1カラムから3カラム、SEXは4カラム
という意味です。
②は旧版と同じようにフォーマット指定をおこなっていますが整数の場合F3.0
等の場合、F3と書いても良くなります。

【例2】 1ケースが複数レコードの際のDATA LISTコマンド
つぎに、1ケースが3レコードから構成されている場合は次のようになります。

```
DATA LIST FILE=INFILE1 RECORDS=3
  /1 NUM 1-3 SEX 4
  /2 AGE 1-2
  /3 T1 TO T5 1-20
```

つぎに、FILE HANDLEと\$PRMFL文の使用方法についての例を示します。

【例3】 FILE HANDLEと\$PRMFL文

```
データの入っているパーマネントファイル (DATA1) のレコー  
ド長が100バイトでファイル定義名をINFILE1とする。

$   JOB      ,A
$   CPROC    SPSSX,,5,220K,-4K,2K
① $   PRMFL  21,R,S,A61234/DATA1
② FILE HANDLE INFILE1/DEVICE=21/LRECL=100
   DATA LIST FILE=INFILE1
     /SEX 1 AGE 5-6 A01 TO A30 8-37 B01 TO B10 85-94
     .
     .
$   ENDJOB
```

- ① \$PRMFL文の“21”はFILE HANDLEのDEVICE=21に対応しています。割当てタイプは読み込みですから“R”となり、アクセスモードはシーケンシャルですから“S”となります。
- ② FILE HANDLE でINFILE1 は、DATA LISTのFILE=INFILE1に対応しています。また、この例ではRECFORM=書式を省略していますがこの場合、RECFORM=Fと同じことです。LRECL=100でレコード長が100バイトであることを指定しています。

2) 自由書式

DATA LIST でキーワード FREEまたは LISTを指定するとカラム位置指定をしなくてもデータ読み込みがおこなえます。FREEとLISTの違いは、FREEでは1レコードに複数のケースが指定できますがLISTでは1ケース1レコードとなります。ただ、“データの見やすさ” からいうと1ケース1レコードにしておくべきだと思います。

【例】FREEの使用法

```
DATA LIST FILE=INFILE1 FREE
      /CNAME(25A) ZONEA ZONEB ZONEC TOTAL*
```

この例は、CNAMEを25桁の文字型変数とし、ZONEAからTOTALが数値型の場合は最後に*を付けます。つぎにデータ例を示しますが、変数間は空白か“,”で区切られていなければなりません。

```
'MATSUDA SEIKO' 7632 9872 6253 23757
'NAKAMORI AKINA' 5862 10156 6532 22550
'KOIZUMI KYOUIKO' 4861 8573 7189 20654
```

3-2. BEGIN DATAによる読み込み

旧版での READ INPUT DATAに代わりBEGIN DATAによってデータの読み込み(80バイト以内)ができます。DATA LISTでは、FILE=を指定しません。

BEGIN DATAは、FILE HANDLEコマンドや \$PRMFL文が不用となります。また、BEGIN DATAとEND DATAコマンドで入力データを挟みます。

【BEGIN DATAコマンド】

```
BEGIN DATA
  入力データ
END DATA
```

【例1】BEGIN DATAの使用例

```
1      8      16カラム
DATA LIST /NUM 1-3 SEX 4 AGE 5-6 T1 TO T5 7-26(1)
      .
      .
BEGIN DATA
$ SELECT A61234/DATA1 .. 利用者番号/ファイル名
END DATA
```

データがパーマネントファイルにある場合、このように\$SELECT文でデータ読み込みがおこなえます。

【例2】複数のファイルからの読み込み

```
.  
.
BEGIN DATA
$   SELECT  A61234/DATA1
$   SELECT  A61234/DATA2
$   SELECT  A61234/TOT-DATA
END DATA
```

【例3】データを直接打つこともできます。

```
BEGIN DATA
00112410.112.314.290.245.2
00203512.434.256.545.731.2
00312134.256.778.456.798.4
.
.
END DATA
```

4. N OF CASESの比較

旧版では N OF CASES によりケース数を指定していましたがX版では不用です。ただし、X版にも同じコマンドがありますが旧版とは意味が違い、ファイルから最初のnケースを取り出すために使います。注意すべきは読むケース数をコントロールするのではなく、作られる実行ファイルのケース数をコントロールしています。

【N OF CASESコマンド】

```
N OF CASES ケース数
```

5. サブファイル処理

旧版では、SUBFILE LIST と RUN SUBFILE によりサブファイル処理をしていましたが、X版ではこれらが廃止され SPLIT FILE により指定します。SPLIT FILE を用いると、実行ファイルを男女別、年齢別などのサブグループに分けてそれぞれのグループで同じ分析を繰り返すことができます。

【SPLIT FILEコマンド】

```
SPLIT FILE BY {変数リスト}
```

【例】男女別，年齢別，に度数分布を求める

```
      .
SORT CASES BY SEX   ... あらかじめケースをソートする
SPLIT FILE BY SEX
FREQUENCIES VARIABLE=T1 TO T5   ... 男女別
SORT CASES BY AGE
SPLIT FILE BY AGE
FREQUENCIES VARIABLE=T1 TO T5   ... 年齢別
SORT CASES BY SEX AGE
SPLIT FILE BY SEX AGE
FREQUENCIES VARIABLE=T1 TO T5   ... 男女別年齢別
```

6. 旧版で作成したシステムファイルについて

システムファイル（実行ファイルを保存したものでデータ定義とデータを併せたファイル）は旧版で作成したのもX版で読むことができます。

【GETコマンド】

```
GET FILE=ファイル定義名 [/KEEP=変数リスト]
      [/DROP=変数リスト] [/RENAME=(旧変数名=新変数名)...][/MAP]
```

ファイル定義名 英文字で8文字以内

/KEEP= で指定した変数を実行ファイルに移す（実行ファイルの順序はここで並べ換えた変数に変わる）

/DROP= で指定した変数は実行ファイルに移さない

/RENAME= 変数名の変更をおこなう

/MAP GETコマンド実行後の実行ファイル内の変数名と順序を表示する

【例】旧版で作成したシステムファイルの読み込み

```
$ JOB ,A
$ CPROC SPSSX,,5,230K,-4K,5K
$ PRMFL 12,R,S,A61234/SYSFILE ... 旧版システムファイル
FILE HANDLE SYSF/DEVICE=12/RECFORM=SYS/LRECL=1024
GET FILE=SYSF
FREQUENCIES VARIABLE=T1 TO T5
```

7. その他のコマンド

- 1) RUN NAME と TASK NAME コマンドは TITLE と SUBTITLE の各コマンドに変更
- 2) LIST CASES は LIST コマンドに変更 そしてPRINT,PRINT EJECT,PRINT SPACE の各コマンドが追加
- 3) OPTIONS と STATISTICS コマンドは変更されていませんが番号が付加されたり，削除されているものがありますから注意が必要
- 4) RAW OUTPUT UNIT コマンドは PROCEDURE OUTPUT コマンドに置き換え
- 5) READ MATRIX コマンドは INPUT MATRIX コマンドに置き換え

8. おわりに

この説明書作成にあたり、ご協力頂いた家本修先生（梅花短期大学）に厚くお礼申し上げます。

9. 参考資料

- 1) 三宅一郎，山本嘉一郎著 新版SPSS[×] I 基礎編（東洋経済新報社）
- 2) 垂水共之，西脇二一，石田千代子，小野寺孝義著
新版 SPSS[×] II 解析編1（東洋経済新報社）
- 3) 家本修 SPSSについて その2-SPSSX入門
センターニュース第55号（1984-11）