

## 公開講演会要旨

# 統計解析ソフトウェア運用システムについて

統計数理研究所 田 村 義 保

(昭和63年11月4日, 統計数理研究所 講堂)

計算機で統計解析を行なう時、優れた解析プログラムが必要であることは言うまでもないことであるが、そのプログラムがどれだけ使い易いかも重要になってくる。パーソナル・コンピュータ用の統計解析プログラムの中には、ユーザインターフェースの良いものもあるようだが、大型計算機用のプログラム(プログラムパッケージ)は市販されているものも含めて、ユーザにかなりの負担をかけるような形になっている。

JCL(ジョブ制御文)というやっかいなものを常に意識しながら書いていかねばならない。これは、普通のユーザにとっては、かなりの苦痛であり、計算機嫌いを生む原因になっているものと思われる。市販されているプログラムパッケージでも、少しは簡単になっているが、やはり、JCLを書く必要がある。

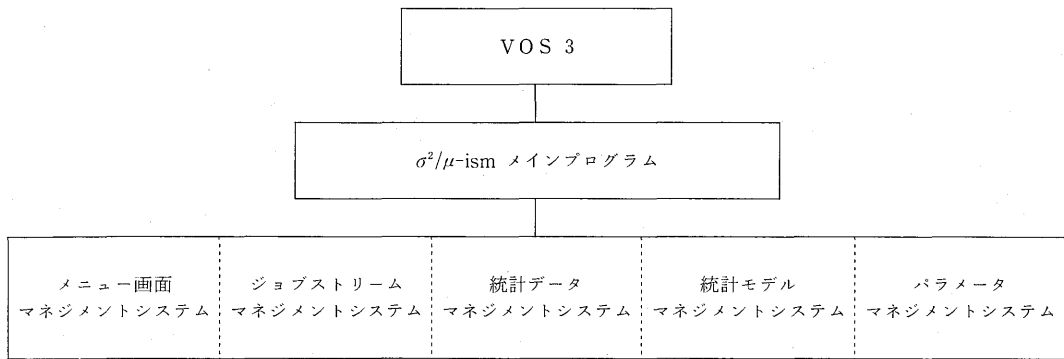
当研究所で開発された最新の統計解析プログラムを大型計算機で使う時に、JCLを意識なくしてすむような環境をユーザに与えるために、本システムを設計・作成するためのプロジェクトが計画された。プロジェクトのメンバーは、大隅 昇、石黒真木夫、中村 隆、桂 康一と田村の5名である。また、システムの詳細設計は(株)日本科学技術研修所に依頼した。

次のような機能を持つシステムを目標として設計・作成を開始した。

- (1) 目的の統計手法を、メニューより選び、ガイド(メニュー)に従って操作して行くだけで、計算を実際に実行するために必要なJCLをシステムが作成する。
- (2) 必要事項を、規則に従ってプログラム中のコメント文に書き込んだ統計解析ソフトウェアを新規にシステムに含めることができる。すなわち、今後開発される統計解析プログラムを整備するためにも役に立つ。
- (3) 統計データ解析で必要とされる複数ステップの統計処理を設定しやすいうように、データ形式の変換機能を有している。
- (4) 将来、統計エキスパートシステムを作る時の参考にするために、詳しい計算記録(ロギング)を残すことができる。
- (5) (3)で述べたことと一部重複するが、市販統計ソフトウェア、所内開発ソフトウェアの中から、必要なものを自由に選択し、組み合わせて使用することができる。

システムの名称は「統計解析ソフトウェア運用システム」(Statistical Software Manager: 略称  $\sigma^2/\mu$ -ism または SSM)とした。また、システム構成は図1のようになっている。

システムの記述言語は: (1) FORTRAN 77, (2) コマンドプロシジャである。また、支援ソフトウェアとして: (1) MODEL, (2) ASPEN を使用している。ここまでの記述を読まれると自明なように、日立製の大型計算機で動作するシステムである。移植性については、あまり重視せずに開発しているようにとられるかもしれないが、プログラムをモジュール化すること

図1.  $\sigma^2/\mu$ -ismシステム構成図

により、機械・OSに依存する部分とそうでない部分を分離しているため、他機種への移植は可能である。

現在のバージョンは以下のような機能を有している。

- (1) モデルとデータ（パラメータ及び統計データ）を指示することによりJCLを自動生成する。
- (2) ジョブ記述ファイル（後述）及びJCLファイルの登録、削除、検索を行なう。
- (3) 作成したJCLをサブミットし、ジョブを実行する。
- (4) 統計モデル記述ファイル（後述）の登録、削除、検索を行なう。
- (5) 統計モデル（ソースプログラム）の画面表示を行なう。
- (6) 統計データ記述ファイル（後述）の登録、削除、検索を行なう。
- (7) 統計データの画面表示を行なう。
- (8) パラメータ記述ファイル（後述）の登録、削除、検索を行なう。
- (9) パラメータの画面表示を行なう。
- (10) フルスクリーン編集対応端末（2020, 560/20等）上でメニュー方式で作動する。
- (11) 市販パッケージを起動する。

システムフローは図2のようになっている。この図に示されているように、各マネジメントシステムは2~3個のファイルと関係している。これらのファイルをオブジェクトとするオブジェクト指向的な考えも一部取り入れて設計している。ここで各ファイルの役割を説明しておく。

- (1) SSM 管理データファイル

$\sigma^2/\mu$ -ismに登録されているファイルに関するメタ情報（ジョブ名、モデル名、データ名、パラメータ名等）を管理するためのファイル。

- (2) モデル記述ファイル

モデル（本システムでは、ある統計解析を行なうための解析手法のことを指している）の内容解析や入・出力情報に関する説明を含む。システムは、モデル記述ファイルの内容を利用して、ユーザがジョブ記述ファイルを作成する時のメニュー画面を出している。モデル記述ファイルの例を図3に示す。

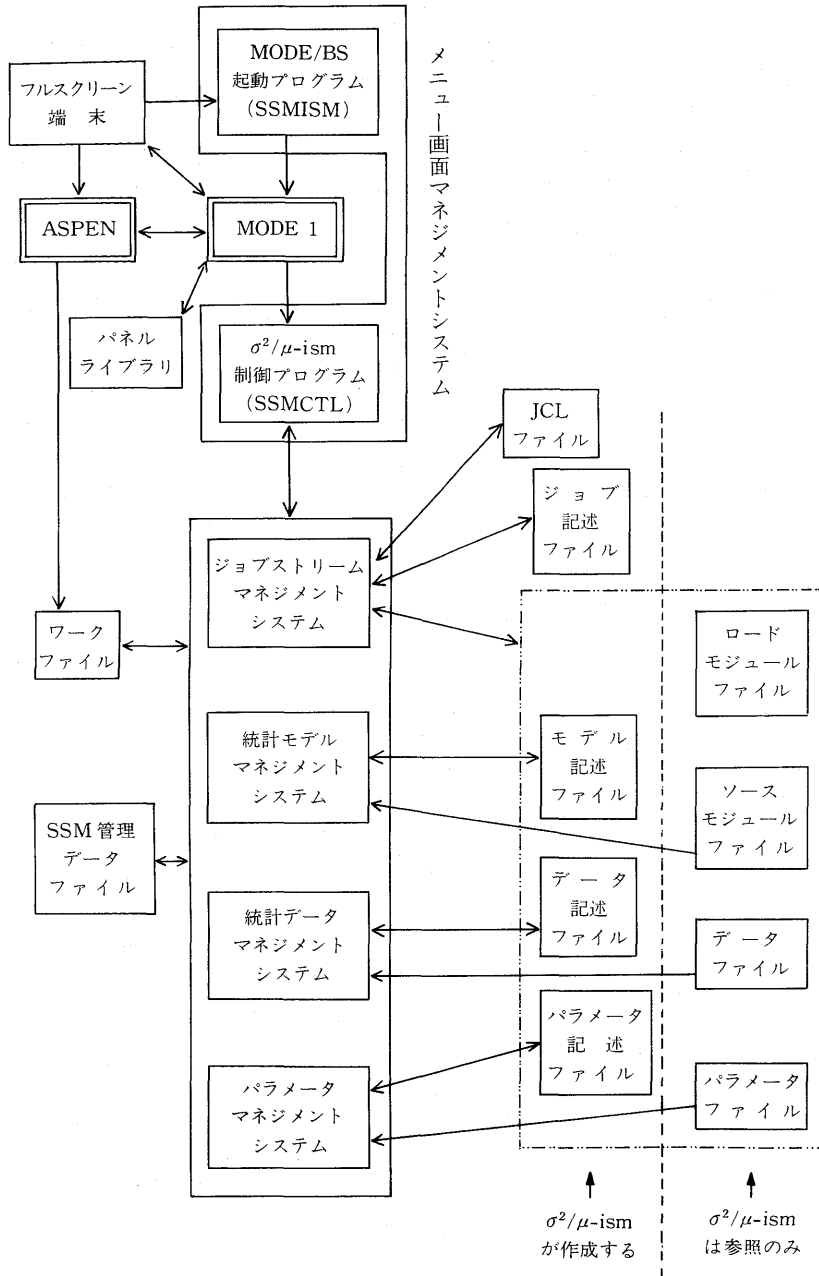


図 2. システムフロー

```

[MODEL NAME]      MULCOR
[SORCE]           TAMURA.TIMSAC.FORT
[LOAD]            TAMURA.TIMSAC.LOAD
[MEMBER NAME]     MULCOR
[ABSTRACT]
MULTIPLE CORRELATION
[DESCRIPTION]
THIS PROGRAM WAS DESIGNED BY H. AKAIKE
THIS PROGRAM WAS PROGRAMED BY E. ARAHATA
THIS PROGRAM WAS MODIFIED BY Y.-H. TAMURA
[AUTHOR]          E. ARAHATA
[REGISTERED DATE] 72/04/20
[MODIFIED DATE]  72/04/20
[INPUT FILE]      1
[UNIT]            12
[KIND]            DATA
[NOTE]            ORIGINAL DATA
[FORMAT]          (415)
[VARIABLE]
N:      LENGTH OF DATA
LAGH:   MAXIMUM LAG
K:      DIMENSION OF THE OBSERVATION VECTOR
ISW:    ROWWISE DATA(1), COLUMNWISE DATA(2)
[FORMAT]          (20A4)
[VARIABLE]
(DFORM(1), I=1,20):FORMAT FOR INPUT DATA
[FORMAT]          (DFORM)
[VARIABLE]
DO I=1,N {(X1(I,11),11=1,K)} DOEND, IF ISW = 1
DO 11=1,K {(X1(I,11),1=1,N)} DOEND, IF ISW = 2
[OUTPUT FILE]     2
[UNIT]            6
[DCB]
[KIND]            PRINTER
[NOTE]            REPORT OF RESULTS A LINE CONTAINS 132 CHARACTERS
[UNIT]            10
[DCB]            (RECFM=FB, LRECL=80, BLKSIZE=3120)
[KIND]            FILE
[NOTE]            INPUT FOR PROGRAM 'FPEC7' OR 'MULSPE'
[JCL]
// EXEC PGM=MULCOR
//STEPLIB DD DSN=TAMURA.TIMSAC.LOAD, DISP=SHR
//*FT12F001 DD UNIT=DISK,
//*DSN=TAMURA.TIMSAC.DATA(FUKU1), DISP=SHR
//*FT10F001 DD UNIT=DISK,
//*DSN=@TAMURA.MULCOR.DATA, DISP=OLD
//
[RELATED PROGRAM] 'FPEC7', 'MULSPE'

```

図3. モデル記述ファイルの例

```

[DATA NAME]          SAMPLE
[DATASET NAME]       #SUPPORT.OKU.TIMSAC.DATA
[MEMBER NAME]        FUKUZ
[AUTHOR]             Y.TAMURA
[REGISTERED DATA]   85/04/01
[MODIFIED DATA]    85/04/01
[ABSTRACT]
  DATA OF AVILITY OF VENTILATION
[DESCRIPTION]
LENGTH OF DATA : 300
MAXIMUM LAG : 25
DIMENSION OF THE OBSERVATION VECTOR : 3
DATA FORMAT : TYPE 2 (COLUMNWISE DATA)
  ___300___25___3___2 : (415)
DFORM : (10F8.3)
      (10F8.3) : (20A4)
ABOVE TWO LINES ARE ADDED TO THE TOP OF DATA
[DATA FORMAT]        (10F8.3)
[DATA TYPE]          9
[COLUMN]              3
[COLUMN NAME]
  NAME1 : VENTILATION
  NAME2 : O2_DENSITY
  NAME3 : CO2_DENSITY
[ROW]                 3
[ROW NAME]
  NAME1 :
[START TIME]
[DATA INCREMENT]    6SEC

```

図4. データ記述ファイルの例

## (3) データ記述ファイル

原データが格納されているデータ・セット名，入力のための制御文（DD文），データのタイプ（1～9に区別されている．データの記述方法は重要な問題であり，現在，統計数理研究所共同研究（63-共研-1）「データ解析の知識ベース構築」の基礎研究」を実施中である．本システムのデータ記述にも，できる限り上記の共同研究の成果を取り入れて行くつもりである）。

## (4) パラメータ記述ファイル

統計プログラムを実行するために設定，入力が必要なパラメータ（実行条件等を設定するための数値）に関する情報を記述するためのファイル．関係する統計モデル名，統計データ名，パラメータの概略説明，実際にパラメータの例が格納されているパラメータ・ファイル名等を含んでいる．ジョブ記述ファイルを作成する際に，この情報が利用され，パラメータの数値例を自動的に取り込むことができる。

## (5) ジョブ記述ファイル

本システムで最も重要なファイルであり、メニュー画面に従って操作していくことにより、自動的に作成される。複数の統計解析を連続して実行するための JCL を簡単に作成できる。入・出力データの割付も簡単に行なえるようになっている。

システムの実際の使用法、画面の動きの詳細を示すには紙面が不十分であるので、本稿はここまでにしたい。詳細を知りたい人は、参考文献を読んでほしい。実際にシステムを利用して頂ければ、より幸いである。

最後に、この場を借りて、システム設計・作成に協力していただいた(株)日本科学技術研修所の SE 諸氏に謝意を表したい。また、本稿を書くために、(株)日本科学技術研修所より提出された詳細設計書と利用マニュアルを参考にさせて頂いたことを述べておく。さらに、本稿の図の一部は利用マニュアルから転載したものであることも述べておく。

## 参 考 文 献

- (株)日本科学技術研修所 (1987). 総合統計解析ソフトウェア, 基本設計書<sup>(4)</sup>.
- (株)日本科学技術研修所 (1988). 総合統計解析ソフトウェア, 詳細設計書.
- (株)日本科学技術研修所 (1988). 総合統計解析ソフトウェア, 利用マニュアル.

## 注

システムの正式名称は統計解析ソフトウェア運用システムであるが、設計依頼時の仮称が総合統計解析ソフトウェアであったため、設計書等の名称が上記のようにになっている。