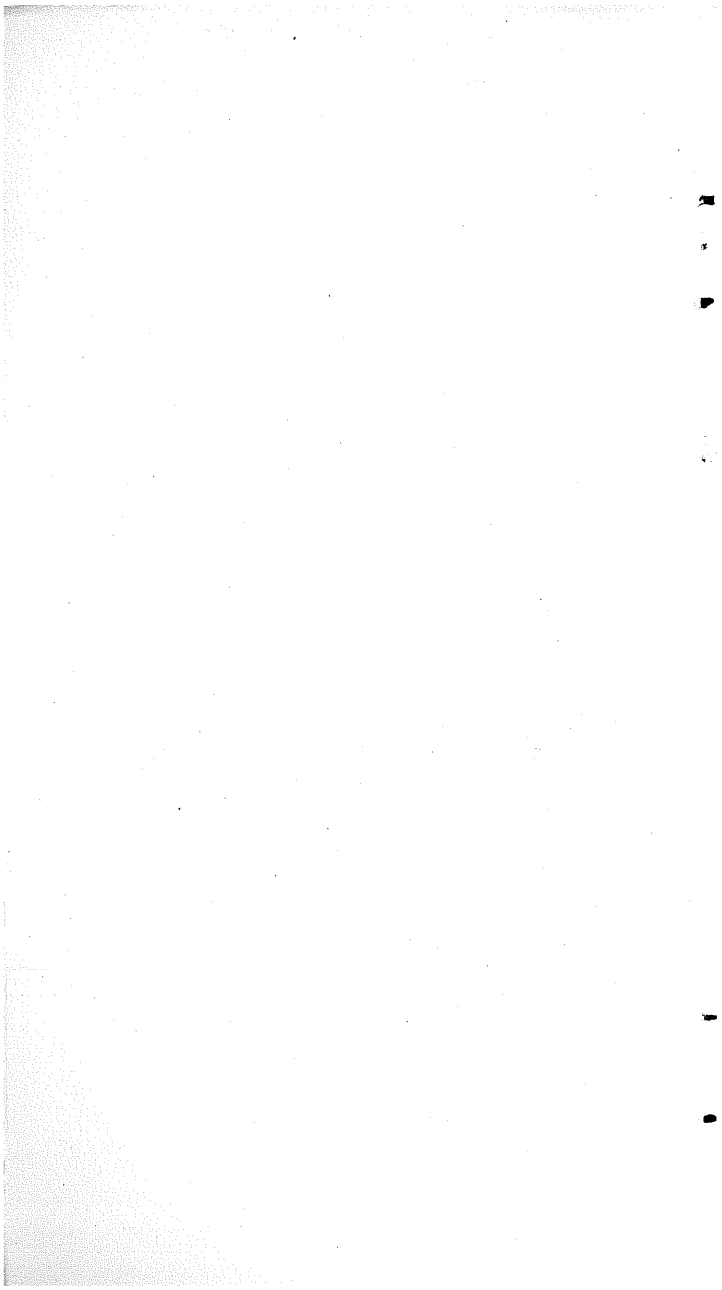




scientific
information
retrieval

**pocket
guide**



SIR
SCIENTIFIC INFORMATION RETRIEVAL

A RESEARCH DATA MANAGEMENT SYSTEM BASED
UPON THE SPSS LANGUAGE

USER'S POCKET GUIDE TO SIR
WITH SECTIONS ON SPSS AND BMDP

AUTHORS

GARY D. ANDERSON
ELI COHEN
WALLY GAZDZIK
BARRY ROBINSON

JANUARY 1979

No written materials describing the SIR system may be produced or distributed without the express written permission of SIR, Inc.

© Copyright 1979 by SIR, Inc. All rights reserved.

SIR - Scientific Information Retrieval
P.O. Box 1404
Evanston, Illinois 60204
(312) 475-2314

INTRODUCTION=====

The SIR POCKET GUIDE has been designed to provide quick referer to the commands and options available to the SIR* user. It also provides a pocket reference to commands and options in both the SPSS and BMDP statistical systems. This guide assumes that the user is familiar with each of these systems and wishes to use this document or for a quick reminder on commands and options. This pocket guide not intended to be a users manual substitute for any of these system. Throughout this guide, items inclosed in [] are optional while [] enclose lists of allowable specifications from which one may chosen.

TABLE OF CONTENTS=====

Contents	Page
-----	----
SIR	
GENERAL RULES.....	1
SYSTEM WIDE COMMANDS.....	6
SYSTEM INITIALIZATION COMMANDS.....	8
SCHEMA CASE DEFINITION COMMANDS.....	9
SCHEMA RECORD DEFINITION COMMANDS.....	11
SCHEMA LISTING COMMANDS.....	16
SCHEMA MODIFICATION.....	16
DATA ENTRY/UPDATA COMMANDS.....	18
BATCH RETRIEVAL COMMANDS.....	20
INTERACTIVE EDITOR COMMANDS.....	26
SPECIAL RETRIEVAL COMMANDS.....	31
SUPPLIED PROCEDURES.....	34
REPORT GENERATOR.....	36
UTILITIES.....	39
EXAMPLES.....	41
EXECUTION OPTIONS.....	45
SPSS	
GENERAL RULES.....	47
CONTROL COMMANDS.....	48
STATISTICAL PROCEDURES.....	49
EXAMPLES.....	65
BMDP	
INTRODUCTION.....	68
CONTROL LANGUAGE.....	68
STATISTICAL PROGRAMS.....	75
EXAMPLES.....	105
INDEX.....	107

SIR GENERAL RULES=====

CHARACTER SET

A-Z, 0-9 may be used for naming variables, files and functions.
0-9, +, -, . and E may be used for numerical representation.

Blanks and commas serve as delimiters. Multiple blanks are ignored except when in string constants enclosed in quotes.

In COMPUTE commands: + add, - subtract, * multiply, / divide
** exponentiation, = is equal to, () grouping delimiters, ; separate multiple expressions appearing in a single COMPUTE command.

In format control: /, *, X, T, I, F, E, A and picture clause edit specifications (see the WRITE command for details).

In logical expressions: LT less than, LE less than or equal to, GT greater than, GE greater than or equal to, NE not equal to, E equal to, with logical connectives AND, OR, NOT.

*SIR, Scientific Information Retrieval; USERS MANUAL; Robinson, Anderson, Cohen and Gazdzik, SIR Inc., Box 1404, Evanston Ill., 60204, 1979.

SIR - GENERAL RULES

VARIABLE NAMES

Variable names may be from 1 to 8 characters in length. The first character must be alphabetic. Names must be composed of letters and/or numbers only. They may contain no special characters or blanks.

There can be a maximum of 4095 variables per record type.

SUMMARY VARIABLE INITIALIZATION DURING RETRIEVAL

The SIR system does not reinitialize the variable values of the summary record to missing values between each new summary record. Consequently, values may be carried forward from previous summary records if they are not overwritten purposely by the user. If variable initialization is required between each summary record, the AUTOSET keyword must be used with the RETRIEVAL command.

FREE FORMAT COMMANDS

Commands	Command names must begin in column 1. Commands must be spelled correctly and completely. Command specifications may begin, following the command name, after one or more common delimiters (blanks or commas).
Continuation lines	Continuation lines can begin in any column after column 1 (i.e., column 1 of continuation lines must be left blank).
Sequencing	If sequence numbers are present in columns 73-80, the NUMBERED YES command must be specified.
Comments	Comment lines require C or COMMENT in column 1. Text then follows one or more common delimiters.

FIXED FORMAT COMMANDS

SPSS Format	As a subset of the above free format specifications, the SPSS control card format may be adhered to (i.e., commands in 1-15 and specifications in 16-72).
-------------	---

DATA TYPES

All non-string variables are handled as real valued in all calculations carried out by the SIR system. Data type options are important for flexibility in inputting data items, for their most efficient storage in the data file and for greater flexibility in data management. Sir supports the following data types:

Type	Examples	Data file storage mode
Integer	-5, 375, 42	Packed into the smallest amount of storage space required for the largest integer that will fit in the number of columns specified for input of this variable.
Scaled integer	4.82, 3.97, 62.2	Converted to an integer for data file storage and reconverted to decimal whenever retrieved from storage.
Real	4.75, 3.03 E-13	One real computer word of storage for each data item.

Date	03/15/75,051273	Stored as the integer number of days since Oct. 15, 1582.
Time	0900, 09/00/32	Stored as the number of seconds since 12 midnight.
String	"JOHN J. JONES"	Stored as a character string variable with length 1 to 12! characters as specified by the STRING LENGTH command. Default length is 20 characters.

 ARITHMETIC OPERATORS AND DELIMITERS

+ Addition	** Exponentiation
- Subtraction	(Open parenthesis
* Multiplication) Closed parenthesis
/ Division	+ Concatenate two strings

RELATIONAL OPERATORS AND LOGICAL CONNECTIVES

EQ-Equals	GE-Is greater than or equal to
NE-Does not equal	AND-Both must be true
LT-Is less than	OR-Either or both must be true
LE-Is less than or equal to	NOT-Logical complement
GT-Is greater than	

MATH FUNCTIONS

ABS-Absolute value	COS-Cosine
ACOS or ARCOS-Arc cosine	EXP-Exponentiation
AINT or TRUNC-Truncation	RND-Rounded integer value
ALOG or LN-Natural logarithm	SIN-Sine
ALOG10 or LG10-Logarithm to base 10	SQRT-Square root
ASIN or ARSIN-Arc sine	TAN-Tangent
ATAN-Arc tangent	TANH-Hyperbolic tangent

DATE AND TIME FUNCTIONS

CDATE (x,f) Convert a date string to a julian integer (the number of days since Oct. 15 1582).

COMPUTE DATE=CDATE ('01/13/73', 'MM/DD/YY')

CTIME (x,f) Convert a time string x to the number of seconds since 12 midnight.

COMPUTE SECONDS=CTIME ('09.33.55', 'HH/MM/SS')

DATET (n1,n2) Returns present date and/or time substring beginning at n1 and ending at n2 from the following 27 character current date/time string.

Columns

Contents

1-3	Day of the week (SUN, MON, TUE, ..., SAT)
4-5	Comma and blank
6-8	Month of the year (JAN, FEB, ..., DEC)
9	Blank
10-11	Day of the month (1 to 31)
12-13	Comma and blank
14-17	Year (e.g., 1977)
18-19	Comma and blank
20-21	Hour (1 to 12)
22	Period
23-24	Minutes (0 to 59)
25	Blank
26-27	AM or PM

SIR - FUNCTIONS

- COMPUTE DATE=DATET (6,17) places the string 'DEC 25, 1978' left justified into the string variable DATE if run on Christmas day 1978.
- JULC (x) Converts the Julian integer x to a 12 character date string of the form 'JAN 14, 1979'.
- WRITE 'THE PATIENT WAS BORN ON' JULC (BIRTHDAT)
- JULN (x,y,z) Returns an integer number of days from Oct. 15, 1582 to the date represented by the integer arguments x, y, and z which represent month, day and year respectively.
- TIME (x) Returns the integer number of seconds from 12 midnight to the time represented by x. The argument x is an integer in the interval (0,240000) where the first two digits are hours, the second two minutes and the last two are seconds.
-

STRING FUNCTIONS

- LEN (x) Returns the integer character length of the string variable x.
- IF (LEN(NAMEA) EQ LEN(NAMEB)) CNT=CNT + 1
- NUMBR (x) Converts the substring x to a binary number.
- SBST (x,n1,n2) Returns the substring of the string variable x beginning at character position n1 and of length n2 characters.
- SRST (x1,x2) Search string x1 for the first occurrence of string x2. Returns the positive integer location where x2 starts if it is found as a separate word. Returns negative integer starting location if x2 is not found as a separate word. Returns zero if x2 is not found.
- IF (SRST(NAME, 'JONES') GT 0) CNT=CNT+1
- This example increments the variable CNT by one if the last name 'JONES' is found as a separate word in the string variable NAME.
-

WITHIN RECORD FUNCTIONS

- AMOD (x1,x2) x1 modulo x2. Returns the remainder of x1/x2.
- CNT (x1,...,xn) Count the number of non-missing values among variables x1,...,xn.
- COMPUTE SUM=CNT(TEST1,TEST2,TEST3)
- FST (x1,...,xn) Return the first non-missing value among x1,...,xn.
- MAX (x1,...,xn) Returns the maximum of the non-missing values among the variables x1,...,xn.

MEAN (x1,...,xn)	Returns the mean of non-missing x's.
MIN (x1,...,xn)	Return the minimum of the non-missing values among the variables x1,...xn.
SRCH (v1,v2,v3)	Search a table of increasing value variables beginning at variable v1 and ending at variable v2 for the location of the value of v3. If v3 matches an entry in the table, the location of the matching entry is returned as a positive integer (counting from the beginning of the table). If no match is found, -n is returned where n is the location of the first value larger than v3. If v3 is greater than the largest value in the table, -n is returned where n is one larger than the number of entries in the table. The indeterminate value is returned if the table v1 to v2 is not in ascending order. AGEGROUP=ABS(SRCH(GROUP1,GROUP10,AGE)) This example classifies AGE into one of ten age groups in the variable AGEGROUP.
STDEV (x1,...,xn)	Returns std. dev. of the non-missing x's.
SUM (x1,...,xn)	Returns the sum of the non-missing x's.

ACROSS RECORD FUNCTIONS

CNTR (x)	Returns the number of non-missing values found for the variable x within the record set.
MAXR (x)	Returns the maximum non-missing value of the variable x within the record set.
MEANR (x)	Returns the mean of the non-missing values of x within the record set.
MINR (x)	Returns the minimum non-missing value of the variable x within the record set.
SUMR (x)	Returns the sum of the non-missing values of the variable x within the record set. SUBTOTAL = SUMR (SYSBP)
STDEVR (x)	Returns the std. dev. of the non-missing x's within the record set.
FSTR (x)	Returns the first non-missing value for the argument x from among values in the set of records processed.
LSTR (x)	Returns the last non-missing value from among the set of records.

SIR - SYSTEM WIDE COMMANDS

SPECIAL FUNCTIONS

COUNT (n)	Returns the count of the number of records presently in record type n for a case.
EXISTS (v)	Determines whether the present value for v has a missing value. Returns 0 if v has a missing or indeterminate value, 1 otherwise. IF (EXISTS(WEIGHT) EQ 1) SUMWT=SUMWT + WEIGHT This example checks to assure the variable WEIGHT has a non-missing value before attempting to add it to SUMWT.
RAND (x)	Returns a pseudo random number in the interval (0,1). The function incorporates a random starting value and will not generate a repeatable sequence.
VARLAB (x)	Given the name of a variable x, this function returns a character string containing the short label of x. If x has no label, a zero length string is returned
VALLAB (x)	Returns a character string containing the value label of the current value of the variable x. A zero length string is returned if no value label exists for the current value. If the current value of x is missing, VALLAB returns a missing label.

----- REQUIRED COMMANDS

Most commands, in each SIR command set, are optional. The following commands, however, must always appear in the specified command set.

command set	commands
System Initialization	FILE NAME, GET FILE, TAPE FILE or DISK RESTORE and PASSWORD
Case Definition	N of CASES, RECS PER CASE, RECTYPE COLS, CASE ID and MAX REC TYPES
Record Definition	RECORD SCHEMA and VARIABLE LIST/INPUT FORMAT or DATA LIST

----- SYSTEM WIDE COMMANDS-----

The commands in this section can appear throughout the SIR command set as indicated.

COMMENT	text
C	text

This command allows placing comment material in the SIR command set for explanatory purposes only. No action results from this command and it may appear as often as necessary throughout the SIR command set.

COMMENT THIS TEXT IS LISTED WITH THE COMMAND SET.

DOCUMENT	text
----------	------

Documentary text which is to be listed with LIST SCHEMA options. Allowed locations are prior to the first RECORD SCHEMA command in the schema case definition command set and between the RECORD

SCHEMA command and the VARIABLE LIST command in each record type definition command set.

DOCUMENT THIS TEXT IS LISTED WITH LIST SCHEMA OPTIONS AND SERVES TO DOCUMENT THE STUDY.

PRINT BACK

{ YES
NO
CONTROL
FORMAT
REPEAT
NOREPEAT }

This command controls the listing of the SIR command set. If no PRINT BACK command is used or if the YES option is used, SIR produces a listing of the user's input commands. In addition, SIR prints a variable-format correspondence table for each INPUT FORMAT or DATA LIST it encounters during a schema definition run. SIR also lists any lines of SIR code generated by DO REPEAT commands.

The NO option suppresses the entire listing. The CONTROL option lists commands only. The FORMAT option lists the format correspondence table. The REPEAT and NOREPEAT control the listing of lines of code generated by DO REPEAT commands.

SPACE n1[,n2]

This command causes n1 lines to be skipped in printing the command set. If n2 is specified, n2 lines must remain on the page after n1 lines are skipped. If there are insufficient lines remaining on the page, page eject occurs.

TASK NAME text

This command resets the subtitle line which is printed at the top of each page of output from a SIR run. Page eject occurs in the SIR output each time this command is encountered and the title, as specified by the RUN NAME command, and the new subtitle are printed before the next command is processed.

TASK NAME THIS LABEL CAN BE UP TO 59 CHARACTERS LONG.

STRING LENGTH n

This command declares the length in characters of every string variable whose initial declaration follows this command until another STRING LENGTH command is encountered. This command can appear as often as necessary to specify different string variable lengths. 20 characters are assigned to each string variable by default and a maximum length of 125 characters is allowed.

STRING LENGTH 36

COMPUTE CHARVAR = "THIS VARIABLE CAN HOLD 36 CHARACTERS".

INPUT MEDIUM

{ CARDS
filename }

Specifies the source of any data to be read into SIR during a run. If CARDS is specified, the data must immediately follow the data read command in the command set itself and be terminated by an end-of-record marker. If a file name is specified, the data will be read from a local file found with that name. More than one INPUT MEDIUM command can occur in a SIR command set to allow input from multiple data files as necessary. This command must appear in the command set before the command which causes the data to be read.

GET FILE
PASSWORD
INPUT MEDIUM
READ INPUT DATA
INPUT MEDIUM
UPDATE REC
FINISH

JSTUDY

DATA1
ACCEPT, ERRFILE = ERRORS1, ALL
DATA2
ACCEPT, ERRFILE = ERRORS2, ALL

SIR - SYSTEM INITIALIZATION COMMANDS

DO REPEAT stand-in variable name = variable list/
 [stand-in variable name = variable list/...]
 one or more replica commands

END REPEAT

This command provides a means of specifying sets of repetitious commands in a shorthand fashion.

```
DO REPEAT                   K=TEST1 TO TEST10/  
IF                         (EXISTS(K) EQ 1)   COUNT = COUNT + 1;  
                              SUM = SUM + K  
END REPEAT
```

This example accumulates a count and a sum of the non-missing test scores among ten test score variables.

FINISH

The last command in the users command set must be FINISH.

```
ERROR LIMIT                 { NONE }  
WARNING LIMIT                n
```

A limit can be specified to the number of error or warning messages respectively generated by SIR. NONE implies no messages. Up to 50 messages are printed by default.

```
ERROR LIMIT                 10
```

SYSTEM INITIALIZATION COMMANDS=====

The commands described in this section may appear only once at the beginning of the user's SIR command set.

EDIT

This command informs SIR that the purpose of this run is to check the syntax of the command set only. No commands are actually executed by an edit run. This command must appear at the very beginning of the command set when used.

```
FILE NAME                   name                                 **
```

To be used with initial SIR creation run only, this command specifies the name to be assigned to the SIR files for this data base. The name can be from 1 to 6 characters long with the first character being alphabetic.

```
GET FILE                    name                                 **
```

This command replaces the FILE NAME command for SIR runs on a previously defined data base. The name must correspond to the name specified in the FILE NAME command used when the data base was originally created.

```
TAPE FILE                   name                                 **
```

When SIR processing is to be done directly from tape, this command is used in place of the GET FILE command. In this mode, cases are processed sequentially directly from tape and not stored to disk. This command assumes the data base has been placed on tape by the TAPE STORE utility. The TAPE FILE command is not implemented for SIR Version 1.1. Implementation of this command is scheduled for SIR version 1.2.

DISK RESTORE name, FILENAME = local file name **
,LOADING=n [/NOREWIND]

If the SIR files reside on tape, this command can be used in place of the GET FILE command to restore the SIR files to disk and assign them to this run. Name is the SIR file name, FILENAME is the tape file name. The LOADING factor, n, is a number between 0 and 1 that indicates roughly how much of the amount of data expected has been input to the file thus far. If not provided, n is assumed to be .5.

**The appropriate one of these four commands is required for every batch SIR run.

NUMBERED { YES }
 { NO }

If this command is present with the YES option, columns 73-80 of the command lines are reserved for line numbering. If not present the NO option is assumed by default and columns 73-80 are available for command specifications.

PAGESIZE { NOEJECT } ,Columns per line
 n

Page control for the listing of the SIR command set is controlled with this command. n specifies the number of lines to occur on each page of output. If NOEJECT is specified, no page ejection is done. Default is 60 lines/page. Cols/line may be within the interval 72-136.

PASSWORD password (Required)

The password is a 1-7 character string which must be provided for every access to the SIR system. The password is set at initial data base creation by the password given with this command on the creation run.

RUN NAME text

This command sets a main title line of up to 59 characters which will be printed at the top of each page of output from SIR run.

SECURITY read password, write password

This command is required using the read/write passwords assigned to the user by the study manager if read and write security has previously been imposed on the data base to be accessed. If no security has been previously imposed, with READ SECURITY and WRITE SECURITY commands, the SECURITY command is unnecessary.

PASSWORD BSTUDY
SECURITY JOHNR,JOHNW

READ SECURITY (security level) password [(security) password]
WRITE SECURITY (security level) password [(security) password]

Assigns security levels (0 lowest, 30 highest) to specified passwords. These passwords may be assigned to users for use in SECURITY commands when accessing SIR. If these commands are not used to establish security on a data base, all users obtain highest security level access to that data base by simply providing the correct password on the PASSWORD command.

SCHEMA CASE DEFINITION COMMANDS=====

The case definition commands are used when defining a new SIR data base. These commands supply information about the nature of the data base as well as general information that applies to each case in the data base.

CASE ID name (Required)

SIR - SCHEMA CASE DEFINITION

Specifies the name of the variable which will identify the case. It must be present with this same name in the variable list for every record type defined later and be consistent in type throughout the data base. It may be of type integer, real or character string. This variable will, by default be included in the CIR to avoid its storage with each record in the data file. The values of this variable must be unique for each case in the SIR data base. If the CASE ID is of type integer or real, it may take on negative or positive values.

CASE ID IDNUMBR

COMMON LIST variable list

Variables named in this command will be included in the common information record (CIR). The CIR is a single record created for each case. During SIR data retrievals, the CIR contents are always available regardless of which record type is being processed within a case.

COMMON LIST AGE,SEX,ETHNIC

MAX KEY SIZE n

Specifies the maximum length in bits that will be required for each SIR index key. This command is necessary only when a record type will be entered after the initial data base creation which will require a larger index key size than any record type entered initially. The size n may be determined by the formula:

SIZE = A + \sum key sizes
where A = \log_2 (# of different record types) + 1
and

Keysize = $2 + \log_2$ (largest possible value) for integer keys

Keysize = 60 for floating point keys

Keysize = $6 * (\max. \# \text{ of characters})$ for string keys

The keys for each record type are the CASE ID, record type number and any sort variables. The largest SIZE calculated among all record types is the value to be used with the MAX KEY SIZE command.

MAX INPUT COLS number

This command is required when one or more of the record types in the data base has more than 80 columns per input line. If it is omitted, the assumed record length is 80 columns.

This command specifies the length of the largest input line of any record type in the data base. Once specified, each input line of each record in the data base is viewed as being this length.

The number specified is always rounded up to the nearest 10 units. The maximum allowable number of input lines for a record type may be calculated by dividing the number specified by this command into 4095 and truncating the result to an integer.

MAX REC COUNT n

Specifies the maximum number of records that will occur in any record type within any case in the SIR data file. A maximum of 1023 is assumed by default.

MAX REC COUNT 300

MAX REC TYPES n (Required)

This command specifies the maximum number of record types that will eventually be specified for this data base. There is essentially no limit to the number of record types that can be defined. This number is normally set somewhat higher than the number of record types assumed for the study initially.

MAX REC TYPES 30

N OF CASES n (Required)

Specifies the maximum number of cases that eventually might be contained in this data base. If the number is not known, a high value should be specified.

N OF CASES 5000

SEQUENCE COLS beginning column[, ending column]

For those record types which will require more than one data card for input, this command specifies the column(s) of the data cards which will contain the card sequencing number. Only record types requiring multiple cards per record are checked in these columns.

SEQUENCE COLS 79,80
SEQUENCE COL 80

RECS PER CASE n (Required)

Specifies the average number of records that are expected for each case in the completed data base. The product of this number and the number specified in the N OF CASES command allows setting up tables for index keys.

RECS PER CASE 150

RECTYPE COLS beginning number[,ending number] (Required)

This command specifies the column(s) in which the record type identification number will be found on each incoming data record.

RECTYPE COLS 6,7

TEMP VAR LIST variable list

This command identifies a set of variables which will be used as scratch or temporary variables in the process of creating other variables later using IF or COMPUTE commands. Variables named in this list will not be stored in the data file. CIR variables, including the CASE ID, cannot appear in this list.

TEMP VAR LIST TEMP1 TO TEMP25

SCHEMA RECORD DEFINITION COMMANDS=====

The commands in this section are used to describe the contents and structure of individual record types. There will be one set of record definition commands for each record type in the study.

During a Schema Definition run, the Record Definition Commands that are used must be placed after the Case Definition Commands. Also the Record Definition Commands of each record type must be grouped together. Different record types, however, do not need to be arranged in sequential order nor do they all need to be defined in a single run.

Within a set of Record Definition Commands, there is a specified ordering of commands that must be followed. The precedence groups among these commands are listed below. Within a group, the commands may appear in any order.

group	command	group	command
1	RECORD SCHEMA	5	COMPUTE IF
2	SORT RECORDS SEQUENCE CHECK REC SECURITY MAX REC COUNT DOCUMENT	6	VALID VALUES VAR RANGES
		7	MISSING VALUES VAR SECURITY

SIR - SCHEMA RECORD DEFINITION

3	VARIABLE LIST-INPUT FORMAT or DATA LIST	8	VAR LABELS VALUE LABELS
4	DATE VAR LIST TIME VAR LIST SCALED VAR LIST INTEGER LIST	8	ACCEPT REC IF REJECT REC IF

ACCEPT REC IF (logical expression)
REJECT REC IF (logical expression)

For data editing at input, logical consistencies which must be true between variable values on incoming records (or values on incoming records with data already in the CIR record) can be specified with these commands. The data record is rejected if the logical expression specified is false for ACCEPT REC IF and rejected if true for REJECT REC IF.

COMPUTE variable=arithmetic expression[;variable=arithmetic expression[;...]]
OR
variable list(index)=arithmetic expression[;...]
OR
variable=variable list(index) [;...]

The COMPUTE command has all of the capabilities of the SPSS COMPUTE and also provides both the ability to include multiple expressions in a single COMPUTE separated by semi-colons and an indexing capability to select a specified member for computation from a list of variables. The index capability is available only for COMPUTE commands used in SIR retrievals.

COMPUTE COUNT=COUNT + 1
COMPUTE AGEGRP=ABS(SRCH(AGE1,AGE5,AGE))
COMPUTE SUMSQ1 TO SUMSQ6(AGEGRP)=SUMSQ TO SUMSQ6(AGEGRP) + SYSBP**2
COMPUTE AGEGRP=ABS(SRCH(AGE1,AGE5,AGE));SUMSQ1 TO SUMSQ6(AGEGRP)=SUMSQ1 TO SUMSQ6(AGEGRP) + SYSBP**2

DATA LIST (# of lines in the record type)/ line number
variable list-columns [(variable type)]
[variable list-columns [(variable type)]]
[/line number...]

This command provides an alternative to the VARIABLE LIST and INPUT FORMAT commands. Variable type may be A or I for character or integer. Floating point variables with fixed decimals are indicated by an integer in place of the variable type indicating the number of places to shift the decimal from its assumed right justified location. If variable type is omitted, SIR assumes the variable is floating point with decimal included or right justified if not included.

DATA LIST (3)/1
CASE ID 1-5 (A), AGE 7-9 (1), BIRTHDATE 20-25(A)

DATE VAR LIST variable list(date format)[/var list(format)]

Variables named in this command will be converted to Julian integers and stored in the data file as the number of days the date represents from Oct 15, 1582.

DATA VAR LIST BIRDATE ('MMIDDIYY')/MARRIED('DDIMMIYYYY')

The I's included in the date format indicates that those columns should be ignored at input (e.g., 'MMIDDIYY' reads 02/14/76 as 021476. A variable included in the DATE VAR LIST must be read as a character string with an A format.

IF (logical expression) computation
[;computation..]

The IF command provides the ability to do conditional COMPUTE commands in that the computation is done only if the logical ex-

pression specified is true. Computation capabilities are identical to the COMPUTE command.

IF (EXISTS(SYSBP) EQ 1) COUNT = COUNT + 1

INPUT FORMAT (format specification)

This command is required immediately following the VARIABLE LIST command and specifies the format for reading the data from the input records to satisfy the variable list. Valid format items X,I,F,A,/,T,(,) and ,. Repetition factors may also be used. Free format input is not supported. The SPSS FIXED keyword is supported.

INPUT FORMAT (I4,1X,A25,5I3,/,T6,2(F5.2,1X))

This example tells SIR to read the first four columns as an integer, skip one column, read 25 columns as an alphabetic string and five variables as 3 column integers, skip to the next input line, tab to column 6 and read two fields five columns wide as decimal numbers with two decimal places to the right. These last two fields are separated by a single blank.

INTEGER LIST variable list (min, max)[/var list (min, max)/...]

Variables included in this command will have an integer identity for data file storage and output. For calculation purposes they will be considered real valued and will be truncated for storage and output listing. Variables specified in the VARIABLE LIST or DATA LIST commands cannot be included in this command since they are already specified as to type. Consequently, this command is for use with variables created by COMPUTE and IF commands.

INTEGER LIST SCORE(1,500)/INDEX(0,20)

MAX REC COUNT n

This command allows the user to specify the maximum number of records that will be found in a particular record type for any case in the data base. The number is set at the MAX REC COUNT specified for the case if this command is not included in the record definition set.

MISSING VALUES variable list (missing value list)[/var list (missing value list)/...]

Up to a maximum of 3 separate values may be specified for each variable to signify missing data. The keyword BLANK may be included to indicate that blanks imply missing. If a missing value is not specified for a given variable, BLANK will be the default missing value. If other values are specified as missing values for a variable, then a blank field on the input data record will be read as zero.

MISSING VALUES SEX TO INCOME(BLANK)/AGE(98,99)

Missing values may also be specified for a string variable. For example,

MISSING VALUES BIRTHDAY("99/99/99"), STATE("ZZ")

If the string variable contains more characters than the missing value, then a match between the left most characters and the missing value will be read as missing.

RECORD SCHEMA number[,name]

This command must appear at the beginning of the record definition command set for each record type. The record type number specified for a record type will normally be expected to appear on each incoming data line in the column(s) specified in the RECTYPE COLS command. To enter data without the record type number on the input lines, see the data input commands. The name is an optional 1-8 character label that can be given to the

SIR - SCHEMA RECORD DEFINITION

record type and used interchangeably with the number when referring to the record type for retrievals.

RECORD SCHEMA 3,LABRECS

REC SECURITY read level, write level

This optional command may be used to impose read and write security levels on the entire record. Only those users with read or write level passwords higher than that specified by this command will be allowed access to this record for reading or writing respectively.

REC SECURITY 0,30

SCALED VAR LIST variable list (power)[/var list (power)/...]

Variables named in this command will be read from the data record and stored in the data file as integers but will be treated as real numbers with the decimal shifted from its assumed right justified location for processing within SIR. The power indicates the power of ten to be multiplied times the integer value when converting it for processing. When such variables are referred to in VAR RANGES or MISSING VALUES commands, they must be referred to by their floating point values.

SCALED VAR LIST WEIGHT (-2)
VAR RANGES WEIGHT (30.00, 60.00)

SEQUENCE CHECK

Causes a check to be made for proper values of all record identifier information on each line of multiple data line record types.

SORT RECORDS variable list

Identifies the list of variables by which the records will be sorted within the record type if multiple records are present for a case. The order in which the names appear, from left to right determines the hierarchy of sorting, major to minor. Sort order is always ascending.

The SORT RECORDS variables must appear in the VARIABLE LIST command of the given record type. They may only be "record variables". They can not be "common", "temporary" or computed variables. They may not appear in a DATE VAR LIST or a TIME VAR LIST. They also can not be "categorical alphabetic variables" whose values are recoded upon entry into the data base.

The SORT RECORDS command is used to establish hierarchical relationships in the data base.

SORT RECORDS VISITDAT, WORKLEVL

TIME VAR LIST variable list (time format)[/var list (time format)/...]

This optional command allows clock time variables to be read as character strings and stored by SIR as the number of seconds from 12 midnight.

TIME VAR LIST JTIME(HHIMMISS),ITIME(HHIMM)

JTIME will read "20.12.53" as 8:12:53 P.M. and store it as 72,773 seconds. ITIME will read "02/30" as 2:30 A.M. and store it as 9000 seconds.

VALID VALUES variable list (value list) [/var list (value list) ...]

For editing discrete valued variables, this command allows a list of valid values to be specified. The user should avoid use of this command for long lists of valid values because of the

excessive amounts of computer time then required for data entry.

VALID VALUES SEX(1,2)/SCORE(1,5,7,9)

Categorical character variables (variables read as character strings but which are to be treated as discrete integer variable within SIR), will be stored in SIR as sequential integers if included in a VALID VALUES command.

VALID VALUES SEX('M','F')
VALUE LABELS SEX ("M") MALE ("F") FEMALE

The value 1 will be stored in the data file for male and the value 2 for female. Notice, however, that when specifying MISSING VALUES or VALUE LABELS for the variable SEX, the character values are referred to.

VALUE LABELS variable list (value1) label1 (value2)label2,.
 [/var list (value1) label1,...]

Up to a 20 character label may be specified for some or all of the categories of each variable.

VALUE LABELS SEX("M")MALE("F")FEMALE/OCCUPTN(1)BLUE COLLAR
 (2)HOUSEWIFE(3)WHITE COLLAR(4)OTHER

During retrievals the value label may be obtained for printing in reports by use of the VALLAB function.

VAR LABELS variable name, label[/variable name, label/...

Variable labels may be as long as desired by the user. Only the first 40 characters or up to the first ';' whichever occurs first will be transmitted to an SPSS save file or the code book listing when the REGULAR listing option is requested. Semicolons in the label also cause skipping to a new line in the DETAILED code book listing and can, therefore, be used for format control in the label.

VAR LABELS VITALCPY, PATIENTS VITAL CAPACITY IN CC;
 THE VITAL CAPACITY IS MEASURED;
 AT THE BEGINNING AND END OF THE EXAM/
 HEARTRT, PATIENTS HEART RATE IN BEATS PER MIN.

During retrievals, a variable label may be obtained by use of the VARLAB function.

VAR RANGES variable list(min, max)[/var list...]

Values falling outside the ranges specified by this command are rejected at data input.

VAR RANGES WEIGHT (30.00, 160.00)/HEIGHT (20,80)

VAR SECURITY variables list(read level, write level)
 [/var list...]

Read and write access levels are assigned as specified from (lowest) to 30 (highest) to those variables included in the list. In order to gain access to a variable, the user's security level-word specified in the SECURITY command must be at least as high as the variable's security level. A level of zero is assigned to each variable by default.

VAR SECURITY NAME, ADDRESS, PHONE, SOCSEC (30,30)

VARIABLE LIST name [name ...] (Required) *

All variables to be defined within this record type must be named by use of either this command or the DATA LIST command. Names consist of 1 to 8 alphanumeric characters and must begin with a letter. The shorthand notation VAR1 to VAR10 will create ten variables VAR1, VAR2, ..., VAR10.

SIR - SCHEMA LISTING

VARIABLE LIST IDNUMBR, NAME, ADDRESS, V4 TO V30

*either a VARIABLE LIST or a DATA LIST command must be included with each schema definition command set.

----- SCHEMA LISTING COMMANDS=====

The schema listing commands provide the user with a flexible set of options for producing reference documentation for the study. The listings made by the LIST SCHEMA command will fit within 8 1/2 x 11 inch page format and are suitable for binding.

LIST SCHEMA keyword list[rt#(variable list) ,rt#...]

The keyword list consists of one or more of the options STRUCTURE, LABELS, REGULAR or DETAILED.

The STRUCTURE option produces a report describing the case and record structure of the SIR data base. Any DOCUMENT commands defined during data schema definition are also listed in this report.

The LABELS option generates a report containing the variable names and short variable labels for the record types and variables specified.

The REGULAR option results in a listing containing variable names, variable type, position in the input record, short variable labels, range or valid values, read and write security levels and missing value designators.

The DETAILED codebook listing contains all of the information in the REGULAR listing with the addition of the full variable label and all value labels.

When variable information is printed, each variable is tagged with a label indicating its origin and location within the record. Labels are of type Vnnn, Cnnn or Tnnn. The nnn is a number indicating the location of the variable within its record. V, C and T represent record variables, CIR variables, and temporary variables, respectively.

A SHORT or LONG keyword may be used to begin the keyword list (except SHORT is not allowed with STRUCTURE). The SHORT keyword suppresses the listing of headings and DOCUMENT output. LONG is default. The SHORT keyword is useful for listings at terminals.

LIST SCHEMA SHORT, LABELS, REGULAR 1,3,7,99

LIST SCHEMA LABELS 1 (SEX TO EDUCATN)

LIST SCHEMA STRUCTURE, LABELS, DETAILED 0,1,2,3,4

A zero included in the record type number argument list produces a separate listing of the CIR record schema.

WRITE SCHEMA FILENAME = filename [/NOREWIND] [RECTYPES = record type list]

Writes a card image code book schema definition command set to the specified file.

LIST STATS

Lists current general information on the data base including: last update, last retrieval, number of accesses, records in each record type, etc.

----- SCHEMA MODIFICATION=====

SIR provides the means of modifying various case and Record definition specifications as required. If data has not yet been entered into the

data base, code book definition sets may be re-entered as many times as necessary with corrections. Once data is in the data base, any change in the schema definition which might affect the data base will cause the record type involved or possibly the entire data base to be "locked" so that it cannot be accessed.

To unlock a locked record type, the following three commands must be executed in order:

1. TAPE STORE
2. PURGE SIR FILE
3. DISK RESTORE

Note: Data Base Administrator security is required to lock or unlock a locked record type.

Often changes in the schema definition are most easily accomplished by re-entry of the complete schema definition set for the entire code book or at least for an entire record type. The SIR Interactive Editor provides several facilities that make such record schema modification rather simple. For example to add two variables RDATE and SDATE to record type 3, the following procedure could be followed:

```
X,WRITE SCHEMA FILENAME = ABC/RECTYPES = 3
```

This writes the entire set of record type 3 schema commands to ABC.

```
READ,ABC
```

This reads the contents of ABC into the SIR Editor. Using the Editor commands, The text is now modified to include RDATE and SDATE.

```
RUN
```

This executes the current text which re-defines record type 3 as desired to include the new variables. If data is in the data base, record type 3 will now be locked and the three step procedure described above will be required to unlock it.

The only commands which may not be changed in an existing SIR data base are CASE ID and MAX INPUT COLS. To change either of these, the user must completely re-enter the data base from raw data. The WRITE SCHEMA and SIR FILE DUMP commands, however, are helpful should this ever become necessary.

Changes in the variable definitions in data bases which already have data in them can have major implications on the data already in the data file. Consequently, the user is advised to read carefully the section on schema modification in the SIR Users Manual and proceed in this area very carefully.

```
MODIFY SCHEMA      rectype number or name[,new name]
```

This command allows the user to modify those portions of an existing record schema that have no effect on the structure of the data file without re-entering the entire schema. The new name field is optional. If a new name is entered, it replaces the old record type name.

The following set of commands are the only ones that are allowed to follow the MODIFY SCHEMA command. They must be entered in the order listed here:

```
SEQUENCE CHECK
REC SECURITY
DOCUMENT
VAR SECURITY
VAR LABELS
VALUE LABELS
ACCEPT REC IF
REJECT REC IF
```

SIR - DATA ENTRY/UPDATE

Changes entered for VALUE LABELS, ACCEPT REC IF and REJECT REC IF in the MODIFY SCHEMA completely replace those presently entered. This means that all value labels must be entered for the variables to be changed and that if the ACCEPT REC IF or REJECT REC IF command is included, the entire set of correct consistency checks must be re-entered.

CHECK { COMPUTE
 VALID
 RANGES
 MISSING
 BOOLEAN }

The purpose of the CHECK command is to provide the user with the means of reverifying existing data against the new specifications entered during schema modification. The various keywords will lead to reverification of data in the data base on all COMPUTE, IF, VALID VALUES, VAR RANGES, MISSING VALUES, ACCEPT REC IF and REJECT REC IF commands in the modified schema respectively. Data is not reverified unless this command is present.

If used, the CHECK command must be placed immediately before the VARIABLE LIST or DATA LIST command in the revised schema set.

----- DATA ENTRY/UPDATE COMMANDS=====

Once the SIR data base has been defined in a Schema Definition run, the user can input, change or delete data in the data base using the commands described in this section.

READ INPUT DATA RECTYPE = number or name
 [/NOSEQ][/NONEW][ACCEPT]
 [/ERRFILE = filename [,NOREWIND][,ALL]]

Data read from the location specified in the latest INPUT MEDIUM command will be added to the data file. Both new and replacement data are allowed. ACCEPT causes records with invalid data items to be stored to the data file with the invalid data items replaced by missing values. The following exceptions will always cause complete record rejection:

1. If the error is in a consistency check
2. If the error is in a record identifier
3. If a line is missing or out of order in a multiple line record

By default all records with error values are rejected before data file entry. The ERRFILE keyword provides a file where copies of rejected records are written. The NOREWIND option prevents ERRFILE from being rewound. ALL causes every record with errors to be written to the error file, even those written to the data file due to the presence of ACCEPT.

The RECTYPE keyword implies that all of the data lines in the data input file are of the same type and do not, therefore, need the record type number on the data lines. (Note - the user must still specify a RECTYPE COLS for the record at schema definition time).

The NOSEQ option instructs SIR to bypass sequence checks for multiple line records. The NONEW keyword prevents the creation of new cases during a data entry run.

READ INPUT DATA ACCEPT ERRFILE = ERRORS, ALL

ADD REC RECTYPE = number or name
 [/NOSEQ][/NONEW][ACCEPT]
 [/ERRFILE = filename [,NOREWIND][,ALL]]

Add new data only, no overwriting of old data allowed.

REPLACE REC RECTYPE = number or name
 [/NOSEQ][ACCEPT]
 [/ERRFILE = filename[,NOREWIND][,ALL]]

Replace old data only, no new data allowed.

```
UPDATE REC          [ERRFILE = filename[,NOREWIND][,ALL]]
                   [/NOCOMP][/NOBOOL]
                   [/MISSCHAR = character]
                   [/RECTYPE = number or name]
                   [/NOSEQ][/ACCEPT]
```

Insert updating data items into existing records. Updating data lines need only contain identifiers and the updated data items in the proper columns. The new value(s) will replace the old value(s).

By default, SIR will perform all COMPUTE commands defined in the schema definition for this record type that are appropriate for the data lines coming in from the input data file. If the NOCOMP keyword is specified, none of the input record COMPUTE commands are executed.

By default, SIR also performs all of the consistency checks defined for this record type on the incoming data. If the NOBOOL keyword is included, ACCEPT REC IF and REJECT REC IF commands are not executed.

The MISSCHAR option allows a character to be specified which will be recognized as a missing value indicator for a data item. If the character specified (blank not allowed) is found in the left most position of a data item field, that data item is set to missing in the data base.

```
MODIFY REC          rt#, case id value, sort id1, sort id2, ...
```

The specified record is made available for modification. COMPUTE and IF statements following the command act on this record.

```
MODIFY REC          3, 1324, 4, 1
COMPUTE             SCORE = 5
IF                  (SEX EQ 1) MARSTAT = 3
```

```
MODIFY REC IF      rt# (logical expression)
```

Makes all records in the specified record type, for which the logical expression is true, available for modification as illustrated in the MODIFY REC command. The logical expression may reference variables in the named record type and the CIR.

```
MODIFY CASE        case id value
```

Allows modification of the CIR record of the specified case.

```
MODIFY CASE IF    (logical expression)
```

Allows modification of all case CIR records for which the logical expression is true. Only the CIR variables may be referenced in the logical expression.

```
DELETE REC         rt#, case id value, sort id1, sort id2,...
```

Purge the identified record from the data file.

```
DELETE REC IF     rt# (logical expression)
```

Purge all records in the identified record type which satisfy the logical expression. The record type and CIR variables may be referenced in the logical expression.

```
DELETE CASES      case id1, [case id2, ...]
```

Purge the specified cases from the data file.

```
DELETE CASE IF    (logical expression)
```

Purge all cases satisfying the stated logical expression. Only CIR variables may be referenced in the logical expression.

SIR - BATCH RETRIEVAL

DELETE CASE IF (CASEID GT 1242 AND CASEID LT 1401)

EVICT REC [RECTYPE = name or number]
[/NOSEQ] /EVICTCIR
[/ERRFILE = filename, NOREWIND]

This command provides a means of removing the data records found on the input data file from the data base. If a set of data records were added to the data base which the user later wishes to remove, this command provides the necessary tool for that purpose.

The data records on the INPUT MEDIUM file must contain (at a minimum) the case id value, the record type number (unless the RECTYPE option is used) and values for all the sort ids, if any are called for. All lines of multiple-line records must be input with complete sequencing information. Any other data on the records is not looked at by SIR. The EVICTCIR tells SIR to eliminate a case completely from the data base (i.e., remove the CIR record) when all of the other records have been evicted.

BATCH RETRIEVAL COMMANDS=====

A SIR retrieval program consists of one or more "blocks" of retrieval commands. Within each block, retrieval operations are performed on records of a single type in the data file or on summary records.

Each block in a retrieval program ends with the beginning of a new block or with an END statement. When END statements are used as block terminators, it is permissible to insert certain statements between blocks.

The general structure of a SIR retrieval program is as follows:

0 or 1	RETRIEVAL block
0 or 1	FOR EACH CASE block
0,1 or more	PROCESS REC blocks
0 or 1	FOR EACH REC blocks
0,1 or more	PROCESS REC blocks
0 or 1	SUMMARY blocks
0,1 or more	procedures (statistics, reports, etc.)
0 or 1	AFTER RETRIEVAL block

The RETRIEVAL block is used for defining data tables, initializing summary variables, printing report preamble information and defining report heading and footing information.

The FOR EACH CASE block is used to limit the number of cases accepted in a retrieval. Without this block, SIR assumes that all the cases in the data base are to be processed.

The PROCESS REC and FOR EACH REC blocks are used to select records from the data base and create summary record variables from these records. The SUMMARY block is used to perform final operations on the variables in the summary record and to make final determinations about whether or not specific summary records will be used by the procedures which follow.

The type of retrieval performed by SIR depends on the presence or absence of the FOR EACH REC block in the retrieval program.

When a FOR EACH REC block is present, a summary record is created for each record of the type specified on the FOR EACH REC command subject to the selection criterion (SELECT REC IF or REJECT REC IF) command included in the block.

If there is no FOR EACH REC block in the retrieval program, a single summary record is created for each case selected from the data base. In this type of retrieval, the summary record for a specific case may be suppressed by using the SELECT CASE IF or REJECT CASE IF command.

The SUMMARY block signals the end of summary variable creation for the current retrieval. All the variables that appear anywhere in the re-

retrieval program become summary variables and appear in the summary record in the order that they are created in the retrieval program.

The AFTER RETRIEVAL block is used for writing report ending information

SIR provides many different kinds of retrieval commands which may appear in the program blocks. The following is a list of these commands grouped by type:

COMMAND TYPE	NAME
Case Control	SELECT CASE IF REJECT CASE IF
Record Selection	SELECT REC IF REJECT REC IF SELECT FIRST EXIT BLOCK IF
Summary Record Selection	SELECT REC IF REJECT REC IF PERFORM PROCS
Summary Variable Initialization	PRESET
Summary Variable Definition	COMPUTE IF MOVE VAR LIST RECODE ASSIGN MISSING VAR LABELS VALUE LABELS VAR RANGES
Report Generation	WRITE WRITE IF HEADING FOOTING PAGE EJECT BLANK LINES

AFTER RETRIEVAL

This statement initiates a retrieval program block which may be located at the end of the retrieval command set after all procedure options. COMPUTE, IF, HEADING, FOOTING, WRITE, WRITE IF, BLANK LINE(S) AND PAGE EJECT commands which follow this command will be executed once at the end of the retrieval.

```
AFTER RETRIEVAL
COMPUTE      AVERAGE = GRTOTAL/GRCOUNT
WRITE        *,'THE AVERAGE PERFORMANCE=',
             AVERAGE (F8.3)
```

```
ASSIGN MISSING  variable list (missing value)[/var list(value)
...]
```

Designates values which are to be assigned to variables created by COMPUTE or IF commands when indeterminate results occur. The assigned value will be used whenever the variable is written out or passed on to procedure. If no ASSIGN MISSING command is used for a variable, the default missing value is BLANK.

```
ASSIGN MISSING  SCORE, AVESCORE(99)/TOTAL(999)
```

```
BLANK LINE
BLANK LINES    n
```

Inserts one or more blank lines in a retrieval output list. Each occurrence of BLANK LINE skips one line. n specified with BLANK LINES gives the number lines to be skipped. If n causes page eject, the remainder of n is ignored.

SIR - BATCH RETRIEVAL

COMPUTE see SCHEMA RECORD DEFINITION COMMANDS
 PERFORM PROCS [(logical expression)[computation[;...]]]

May appear within a SUMMARY block. Normal output of the summary record is suspended and a single copy of the summary record is written out for each PERFORM PROCS command encountered (if a logical expression is present, it must first evaluate to true). Makes possible output of multiple copies of a summary record with computations controlling the content of each one.

```
SUMMARY
DO REPEAT      I = T1 TO T5/
PERFORM PROCS (EXIST(I) EQ 1) DRUG = I
END REPEAT
END
```

END

Terminates a retrieval block. May be omitted in cases where the end of the retrieval block is obvious. Required if the user desires to place commands between two blocks.

EXIT BLOCK IF (logical expression)

This command may only appear in a PROCESS REC or FOR EACH REC block. It causes termination of record processing within a retrieval block when the logical expression is true. Execution then jumps out of the block if the command is in a PROCESS REC block and jumps to the next case if the command is in a FOR EACH REC block.

```
PROCESS REC      4
SELECT REC IF   (SEX EQ 2)
EXIT BLOCK IF   (VISTDATE GT JULN(01,01,77))
MOVE VAR LIST   BLOODPRS TO VITALCPY
END
```

FOOTING output specification

The output specification will be written at the bottom of each page of output from a retrieval prior to page eject. The output specification is as described in the WRITE command.

FOOTING 100X, 'PAGE.',PAGE

```
FOR EACH CASE  one
                only { COUNT = total[,increment[,start]]
                     LIST = case id1[,case id2,...,[case idn
                     THRU case idm]...]
                     SAMPLE = fraction [,seed]
```

Located after the RETRIEVAL command and prior to any other retrieval program blocks, this command provides information about which (or how many) cases are to be processed by this command set.

```
FOR EACH CASE  COUNT = 25
FOR EACH CASE  COUNT = 300, 1, 3050
FOR EACH CASE  LIST = 931, 1122, 2400 THRU 2851, 5001
FOR EACH CASE  SAMPLE = .01, 1785
```

```
FOR EACH REC rt# [[ WITH
                   UNTIL
                   THRU
                   FROM
                   AFTER ] (sort id list) ] [ { UNTIL
                                                THRU } (sort id list) ] ]
```

Signifies the beginning of a master processing block and initiates reading of each record within the specified record type. A summary record is produced for each record in the record type which is selected by the SELECT REC IF command, if any, present in this processing block. Using the specified keywords and sort id list, reading within the record type can be started and stopped at specified locations based upon values of the sort

variables. The sort variables must be specified in the sort id list in the order they were defined at schema definition time by the SORT RECORDS commands.

```
FOR EACH REC      3
FOR EACH REC      3, WITH(3,2)
FOR EACH REC      3, FROM(VISTNUM,1)THRU(VISTNUM,5)
```

HEADING output specification

The output specification list will be written at the top of each page after each page eject. The output specification is described in the WRITE command.

HEADING 'CASE REPORT' ,80X, 'PAGE.',PAGE

IF see SCHEMA RECORD DEFINITION COMMANDS

IF (SEX EQ 2 AND AGE GE 21) INDEX = INDEX+1

MOVE VAR LIST variable list

Values for the variables in the variable list are moved directly from the record being processed into the summary record. Variable labels, value labels, missing values, etc. in effect or these variables are made available for SIR procedures which follow. This command may appear only in a PROCESS REC or FOR EACH REC block.

MOVE VAR LIST AGE, SEX, EDUCATN

PAGE EJECT [n]

Causes page ejection in the retrieval output listing. If n is specified, page eject will occur only if fewer than n lines remain on the page.

PRESET variable list(assignment value list)[/var list (values)/...]

Initializes variables to specified values. The frequency with which the variables are preset is determined by the location PRESET command in the retrieval set:

- If located in the RETRIEVAL block, the variable listed are set at the beginning of the retrieval and unchanged by SIR throughout the remainder of the retrieval.
- If located in a FOR EACH CASE block, the variables listed are reinitialized at the beginning of each new case.
- If in a PROCESS REC or FOR EACH REC block, the named variables are re-initialized once each time the process block begins execution.
- If in a SUMMARY block, the variables initialized at the beginning of the retrieval are re-initialized at the beginning of each new summary record.

PRESET TOTAL1 TO TOTAL10 (0)/SUM1, INDEX (0,1)

The PRESET command must be located immediately after the block heading. If the variable list is longer than the assignment value list, the assignment value list is repeated until the variable list is satisfied. MISSING is a valid member of the assignment value list.

PROCESS REC {rt# name} [{ WITH UNTIL THRU FROM AFTER } (sort id list)] [{ UNTIL THRU } (sort id list)]

Signifies the beginning of a processing block which initiates the reading of all records in the specified record type. If located before a FOR EACH REC command, or if none is present, the PROCESS

SIR - BATCH RETRIEVAL

REC block is executed once for each case. If located after a FOR EACH REC block, it is executed once for each summary record created within that case.

```

PROCESS REC      3
SELECT REC IF   (AGE GE 21 AND AGE LE 40)
MOVE VAR LIST   EDUCATE TO INCOME
END
PROCESS REC      3,WITH(5,1)
.
.
PROCESS REC      2,AFTER(FAMMEMBR,VISTDATE,1)
                UNTIL(FAMMEMBR,VISTDAT4,4)
.
.

```

```

RECODE          var list(value list = new value)...(value list=
                value) [var list...] if the recode does not
                change the recoded variable type or

                new variable=old variable(value list=new value)
                (value list=new value) [/new variable=old vari-
                able (CONVERT)...] if the recode results in a
                change in variable type.

```

This command allows specifying new variable values and collapsing multiple values into single values. Since direct recoding of character to numeric values would be attempting to change a variable from character to numeric, it is necessary to recode character values into a new variable which is numeric in type as indicated by the second form of the command above.

```

RECODE          INCOME (0 THRU 5000=1) (5000 THRU 10000=2)
                (10000 THRU HIGHEST=3)

RECODE          NEWCLAS=SOCCLAS('LOW'=1) ('MED'=2) ('HI'=3)

REJECT CASE IF (logical expression)
SELECT CASE IF (logical expression)

```

These commands may appear in any SIR retrieval block except RETRIEVAL and AFTER RETRIEVAL and they may be placed between blocks. If the SELECT (REJECT) CASE IF logical expression is false (true) then execution for that case stops and processing of the next case begins.

```

REJECT REC IF   (logical expression)

```

Located within a PROCESS REC or a FOR EACH REC block, this command determines which records from the record type being read are to be processed by the commands within the block. Records are rejected if the logical expression is true.

```

REJECT REC IF   (SEX EQ 2)

```

```

RETRIEVAL       [AUTOSSET[,UPDATE]]

```

This command signals the beginning of a retrieval block and must be the first command in the retrieval command set. The AUTOSSET keyword causes SIR to initialize the values of all summary variables in the summary record, that are not named in a PRESET command, to missing at the beginning of the retrieval and each time before a new summary record is begun. It AUTOSSET is not used, SIR does no automatic summary variable initialization and values from previous summary records will carry forward unless initialized by the user. The UPDATE keyword allows modification of data in the data base by the retrieval.

```

SELECT REC IF   (logical expression)

```

Located within a PROCESS REC or a FOR EACH REC block, this command determines which records from the record type being read are to be processed by the commands within the block. Records are accepted if the logical expression is true.

```

PROCESS REC          7
SELECT REC IF      (SEX EQ 1 AND EXISTS (SIZE) EQ 1)
COMPUTE           SUM1 = SUM1 + SIZE ; CNT = CNT + 1
END

```

SELECT FIRST (logical expression)

Located in place of a SELECT REC IF, this command causes selection of only the first record for which the logical expression is true.

```

SELECT FIRST      (BLOODPRS GT 140)
SUMMARY          [ { INCLUDE }
                  { EXCLUDE } = variable list ]

```

This command signifies the beginning of a processing block which completes and passes the summary record to the procedures which follow. Final transformations may be performed on variables within the summary record or additional variables may be created by including necessary commands in this block. By default, unless multiple PERFORM PROC commands are included, a single copy of the present summary record is passed on to the procedures which follow each time this block is executed. By default, the summary record will contain every variable defined or moved by the retrieval program in the order defined. The EXCLUDE option allows indicating variables not to be included in the summary records passed to the procedures. INCLUDE specifies the set which is to be included and in what order.

```

SUMMARY          EXCLUDE = TEMP1 TO TEMP10, AVECOST
COMPUTE          AVECOST = COSTOTL/NVISTS
WRITE           'THE AVERAGE COST IS',AVECOST(F8.2)
END

```

VALUE LABELS variable list(value 1) label 1(value 2)label 2.
[/variable list (value 1)...]

This command permits the user to specify labels for individual values of one or more summary variables. The command can be placed anywhere in the retrieval command set.

VAR LABELS variable 1, label 1 [/variable 2, label 2...]

This command allows the user to define variable labels for computed summary variables.

VAR RANGES variable list (low,high) [/variable list...]

This command allows the user to define a legal range for computed summary variables. Values falling outside the specified range will be assigned missing values.

```

WRITE           output specification
WRITE IF       (logical expression) output specification

```

These commands provide a means of writing information to the output list device during a retrieval. Values may be printed in free format or under either format or edit specifications:

Format specifications:

- Iw - Print the whole portion of an integer or floating point variable in a field w columns wide with no decimal point. *Fill if unable to fit.
- Fw.d-Print an integer or floating point variable right justified in a field w columns wide with d digits to the right of the decimal point. *Fill if unable to fit.
- Aw - Print a character string left justified within a field w columns wide. Right most characters dropped if unable to fit.

SIR - EDITOR

Edit Specification:

```

g blank fill to the left
Z Zero fill to the left
. decimal point in a column (one per edit)
, comma in a column (more than one allowed)
$ floating dollar sign (one per edit)
* asterisk protected field (e.g. for check printing)
  
```

The following example illustrate some uses of the edit specification:

Value	Edit Specification	Output
136.38	('99999.9')	bb136.4
136.38	('ZZZZZ.9')	00136.4
136.38	('99999.9')	b\$136.4
136.38	('99999.9')	\$0136.4
150	('ZZZZZ')	00150
9742.50	('999,999.99')	b\$9,742.50
136.38	('999,999.99')	bbb\$136.38
136.38	('***,***.99')	****136.38
9742.50	('***,***.99')	**9,742.50

Line, page and column control:

each / implies carriage return line feed.

*at the start of the list causes page eject.

nX causes spacing of n columns

nT causes tabbing to column n

Special functions allowed:

DATE prints present date.

TIME prints present time.

PAGE prints present page of output.

```

WRITE 40T,          'THIS REPORT WAS PRINTED ON', DATE, 'AT', TIME
WRITE IF           (SEX EQ 1) 'MALE RESULTS = ' MEN(F8.2)
WRITE              'THE ANSWERS ARE', 3X, VAR5(F8.2), 3X, VAR6(F8.2)
WRITE IF           (SEX EQ 2) 'FEMALE RESULTS = ',WMEN('99999.99')
  
```

INTERACTIVE EDITOR COMMANDS=====

The SIR Interactive Subsystem provides the facilities for performing interactive information retrieval from the user's data base. During a SIR Interactive run the user can:

- input a set of SIR retrieval commands and correct and edit them with the interactive Editor,
- execute the retrieval commands receiving the results at the terminal or directing them to local files, or offline printer,
- save the retrieval commands in the SIR data base stored procedures file for later use,
- execute previously defined sets of commands from the stored procedures file,
- specify run-time parameters for a given retrieval,
- have the SIR retrieval program interact with the user by prompting for and accepting input during execution,
- change selected data base variables using COMPUTE and IF statements (provided the UPDATE option is included with the RETRIEVAL command).

The SIR Interactive Subsystem is invoked by executing SIR with the following command:

SIR, IA.

This command indicates that the SIR interactive version is to be executed. In addition to IA, the following parameters may be specified in the SIR command:

FN = SIR file name
PW = Password
RS = read security
WS = write security
PS = Page size (lines)
WD = page width (cols.)
EP = named procedure to be executed immediately

Required parameters will be prompted for by SIR if not provided in the SIR command. PS and WD will default to 60 and 136 respectively if not included in the command. The user may wish to set these parameters to a size appropriate for his terminal. If PS and WD are included in the SIR command without values specified, they default to 24 and 80 respectively. See the SIR EXECUTION OPTIONS section for more details and information on the ENTER C/H/N/S/BLANK CR message.

In the description of commands that follow, the comma and blank can be interchanged, i.e. either one can be used to separate the fields of an Editor command. No other characters are legal as field separators. Conversely, any special characters can be used as text delimiters (on the SEARCH and REPLACE commands) except for comma and blank.

Where line #s are referenced in a command, the * symbol may be used in place of a number to designate the last line in the present text. For example;

L,*

results in the last line of the current text file in the SIR Editor being listed.

ACCEPT ,local file name (AC)

Copies a local numbered text file, created previously by the editor or WRITE command, into the editor. RACCEPT rewinds the local file before copying. RRACCEPT rewinds before and after copying.

ACCEPT, TEMPFIL

ADD [,first line#[,increment]] (AD)

This command informs the editor that new text is to be entered at the terminal. SIR will then prompt with line #s starting at the specified first line # and incrementing thereafter by the increment specified. Increment is 10 by default. If no line # is specified, SIR starts at 10 if the editor presently contains no text or the last existing line # plus 10 if there already is text. An '=' in response to a prompt informs the editor to terminate the add mode. Note: the editor will overwrite and interleave with existing lines.

ADD, 10, 5

Single lines of text may also be entered by typing a line number, a period and the desired line of text. For example, to insert a line between 70 and 80 of the current text, the user could type:

>75.SELECT REC IF (WID*HGT GE 25)

If the user specifies a line number that already exists in the text, the new line replaces the existing one.

CLEAR (CL)

Clears all current text from the editor work area.

SIR - EDITOR

COPY ,source line, destination line (CO)
 COPY ,source begin, source end, destination
 COPY ,source begin, source end, destination, incr

This command produces copies of text from one location at another location. The original text is left undisturbed.

```
>COPY ,110, 370
>COPY ,270, 370, *
>COPY ,50, 290, 4000, 3
```

In the first example, line 110 is copied into line 370. In the second example, lines 270 through 370 are copied after the the present end of the text. (when the * is specified, the destination line becomes the present last line plus the increment). In the third example, lines 50 through 290 are copied to lines 4000, 4003, 4006, etc.

CWRITE ,local file name (CW)

Saves the current text without line numbers. A file written by this command is in card image format. The command RRCWRITE leaves the resulting file in rewound form.

DELETE ,line # (D)
 DELETE ,first line #, last line #

Removes the designated line(s) from the editor text. When the second form of this command is used, all lines in the interval are delete inclusive.

```
DELETE, 214
DEL, 181, 185
```

FINISH (F)
 END (E)
 QUIT (Q)

Terminates execution of the SIR Query Subsystem session.

GET ,procedure name (G)

Causes the text of the named procedure to be copied from the user's SIR procedure file into the editor work area. Any text presently in the work area is first cleared.

GET, PATLIST

HELP (H)
 HELP , command name

Provides the user with helpful information on SIR Query commands. The first form lists a short description of every command while the second form lists a description of the specified command.

LIST (L)
 LIST ,line #
 LIST ,first line #, last line #
 LIST ,NAMES

Lists all of the current editor work area text, a specified line of text, an interval of lines or a list of the names of the defined procedures currently in the user's SIR procedure file.

```
LIST ,215
LIST ,9,*
L ,NAMES
```

MOVE ,source line, destination line (M)
 MOVE ,source begin, source end, destination
 MOVE ,source begin, source end, destination, incr

This command moves one or more lines of text from one part of the users workspace to another. The original or source lines of text

are deleted. The arguments used with this command are identical to those used with the COPY command.

OUTPUT ,local file name (O)

Placed just before a RUN command, this command diverts all output from the execution to the named local file. Output defaults to the terminal again as soon as the execution resulting from the RUN command is completed.

PAGE ,number (PA)

This command sets the page size for output at the terminal where the number given is the number of lines to be printed per page. The page size can be changed whenever the user wishes. Paging can be turned off completely by typing zero for the page size.

After a page has been printed, the system types:

ENTER C/H/N/S/BLANK CR

A response of 'C' continues this listing without further paging 'H' prints help on the command while 'N' or 'S' both terminate further printing. 'N' completes the present execution without further output printing while 'S' terminates all processing immediately and returns control to the editor. A blank space followed by a carriage return tells the system to continue printing the next page.

PREAD ,file name (PR)

This command reads a file containing one or more SIR procedures as previously written out by the PWRITE command. RRPW rewinds before and after reading.

PURGE ,procedure name list (PU)

Deletes one or more procedures from the user's procedure file.

PURGE, PATLIST ,SORTST, REPORT1

PWRITE ,filename (PW)

This command creates an external, non-SIR file containing a copy of all the procedures presently in the procedure file of the current data base. RPW rewinds the file before writing on it while RRPW rewinds the file both before and after writing on it.

READ ,local file name (REA)

Copies a local card image unnumbered file into the editor work area. Line numbers are automatically supplied beginning at 10 and incremented by 10. If text already exists, the new lines are appended at the end and incremented by 10.

READ, CARDFIL

REPLACE ,d text1 d text2 d (REP)

REPLACE ,d text1 d text2 d,line #

REPLACE ,d text1 d text2 d, first line #, last line #

Replaces text1 by text2 in all lines, in a specified line only, or within a specified interval of lines. Any special character can be used as delimiters 'd'.

The first form of the REPLACE command, which replaces text in all lines, should be used carefully. Since each line of text in which a change is made is listed at the terminal, it is suggested that the page size be set to 1 when using this command. In this way, the results of the REPLACE command can be observed in the first line modified before instructing the system to continue with changes in all of the text lines. (see the PAGE command.)

SIR - EDITOR

```
REPLACE      *JONIS*JONES*
REP 'THIS IS'THESE ARE',39
REP/HALP/HELP/,10,400
```

The three forms of the REPLACE command described above can also contain column anchors which make the search for the text1 string more efficient. The general forms are:

```
REPLACE      ,C1 d text1 d text2 d
REPLACE      ,C1,C2 d text1 d text2 d
```

In the first form, the replacement is only preformed on those lines which contain text1 starting in column C1. In the second form, the replacement occurs if text1 begins anywhere between C1 and C2 inclusive.

```
RESEQUENCE  [,starting line #[, increment]] (RES)
```

Resequences all the line numbers. If given without an argument, the text is resequenced by increments of 10 starting at 10.

```
RESEQUENCE      385,10
```

```
RETURN      ,local file name (RET)
```

Returns the named local file.

```
REWIND      ,local file name (REW)
```

Rewinds the named local file.

```
RUN (RU)
```

```
RUN      ,(parameter ,parameter ,...)
```

```
RUN      ,procedure name
```

```
RUN      ,procedure name (parameter ,parameter ,...)
```

Causes execution of current editor work area text or a named procedure from the SIR procedure file. Parameters may be supplied as required by the procedure. If a procedure is run from the procedure file, present work area text is first cleared. Use of the OUTPUT command prior to the RUN command will divert the output from the execution to a local file. By default, output from the execution returns directly to the terminal. See the CALL command for details on the parameter list.

```
RUN      ,(1,32)
RUN,PATLIST (1292,1840)
RUN
```

```
SAVE      ,procedure name[,OK] (SA)
```

Specifies that the current text is to be stored into the SIR procedure file from where it can later be retrieved for further modification or run as an SIR user defined procedure. The procedure name can be any name up to 7 characters in length starting with an alphabetic character. The OK keyword, if present, allows a present file with the same name to be overwritten. If OK is not present, the system will prompt for it before overwriting an existing file.

```
SAVE ,PATLIST
```

```
SEARCH      ,d text d (SE)
```

```
SEARCH      ,d text d, line #
```

```
SEARCH      ,d text d, first line #, last line #
```

Searches for the specified text in all of the text, in the line indicated or within the interval given. Any special character can be used as delimiter 'd' except blank and commas. Each time a line is found containing the text specified, the line is printed at the terminal.

```
SEARCH, *JONIS*
SE 'THIS IS',45,60
```

Column anchors can be used with any of the three forms of the SEARCH command. The general forms are:

```
SEARCH ,C1 d text d
SEARCH ,C1, C2 d text d
```

If the first form of SEARCH is used, SIR will only list those lines in which the text starts in column C1. If the second form is used, SIR will list the lines if they contain the text string anywhere between columns C1 and C2, inclusive.

```
TAB ,column 1 [,column 2,...] (T)
```

Defines one or more tab columns to be used when lines of text are added. The tab character is the ampersand '&'. Whenever the tab character appears in a line of text being input, the remaining text in the line is placed starting at the next tab stop. The default tab stop is 16.

```
TAB ,16, 28, 40
```

```
WRITE ,local file name (W)
```

Saves the current numbered lines of text as a local temporary file. It can be read in again with the ACCEPT command. The command RRWRITE produces a rewind file.

```
X ,command
```

The X command is designed primarily to perform the SIR utility routines and any SIR stand alone command that can be typed in a single line. The command is executed immediately without disturbing the text presently in the users workspace.

```
X, LIST SCHEMA LABELS 3
X, LIST STATS
```

```
EP control card parameter
```

The EP control card parameter allows immediate execution of a specified procedure upon entry into SIR. Using this parameter in the SIR control card as part of a computer system command file allows a user to execute SIR and enter a specified procedure by merely typing a single word. For example, the following control card would cause SIR to immediately execute the procedure NEWP from the users procedure file:

```
SIR, IA, PS, WD, EP=NEWP, FN=INFA, SP.
PW=BABS, RS=RBABS, WS=WBABS.
```

```
$ ,command.
```

The \$ command allows operating system commands to be executed from within SIR.

```
$, FILES.
```

Note: the \$ command statement must be terminated by a period.

SPECIAL RETRIEVAL COMMANDS=====

SIR provides the user with several extensions to the batch retrieval capabilities described under the section entitled BATCH RETRIEVAL COMMANDS. These commands are especially designed for use with the SIR interactive subsystem but, where appropriate, may also be used in batch. They provide:

- a CALL facility for nesting procedures within other procedures together with a full parameter substitution (macro) capability,
- two functions for providing interactive prompting for input from executing procedures and extensions to other retrieval commands to allow full use of these functions,

SIR - SPECIAL RETRIEVAL COMMANDS

- commands to allow direct creation of a single case or a single record from within a retrieval.

CALL procedure name[(parameter , parameter ...)]

Allows calling a stored procedure from within another procedure. The text of the called procedure is placed within the text of the calling procedure. If a parameter list is included, the called procedure must be coded with a number in brackets in each location the respective parameter is to be substituted (numbered from the left in the parameter list).

CALL CASENAME (I, CASE ID)

The rules for forming the parameter list in the RUN and CALL commands are:

- 1) The parameter list consists of numbers or character strings separated by commas.
- 2) The entire parameter list must be enclosed in parenthesis.
- 3) A list of items will be treated as a single parameter if it is enclosed in parenthesis. For example; (15, EXAM1, (A1,A2,A3)) has 3 parameters; 15, EXAM1 and A1, A2, A3.
- 4) If a parameter is enclosed in quotes, the quotes are considered as part of the parameter.
- 5) Quoted strings containing commas must be in parenthesis to be considered a single parameter.
- 6) Null parameters may be transmitted by two commas in succession or by empty left and right parentheses.
- 7) A parameter list cannot extend beyond one line and it cannot result in the line into which a parameter is substituted extending beyond its present maximum line length.

CASE IS case id

Causes the case with the specified id value to be created or retrieved in the data base. UPDATE is required on the RETRIEVAL command if a case is to be created.

```
RETRIEVAL UPDATE
COMPUTE ID = NREAD ('ENTER CASE NUMBER')
CASE IS ID
COMPUTE SEX = NREAD ('ENTER SEX(M=1 F=2)')
```

DELETE CASE

Deletes the case presently retrieved. All records belonging to the case, including the CIR, are deleted. The user must have security clearance higher than that of any record or variable in the case and have used the UPDATE option in the RETRIEVAL command.

DELETE RECORD

Deletes the record currently retrieved from the data base. The user must have used the RETRIEVAL UPDATE command and have higher security clearance than that of the record or any variable in it.

```
RETRIEVAL UPDATE
PROCESS REC 2
SELECT REC IF (BGNDATE LT JULN(4,1,73))
DELETE RECORD
FINISH
```

EXIT

Causes termination of the procedure presently executing and returns control to the interactive editor.

```
IF (logical expression) {
    computation
    JUMP label
    EXIT
    WRITE
}
```

[DELETE CASE
DELETE RECORD]

Extends the batch IF command to allow any of the listed arguments in addition to computation. Multiple computations, separated by ';' are allowed as defined previously for the IF command.

```
IF(COUNT EQ 1) ANSWER-SPREAD('DO YOU WANT IT LISTED')
IF(BIRTHDATE EQ DATE) JUMP EARLY
IF(COUNT GT 50) EXIT
```

JUMP label

Causes execution to branch to the command with the specified label name. Labels can be applied to any command (other than a DO REPEAT or DO END command). A label is a 1-8 alphanumeric character name terminated by the ':' symbol which precedes the command being labeled. A label can appear in a line by itself or the convention LABEL : CONTINUE may be used.

```
LABEL1 : COMPUTE SEX = SREAD ('ENTER SEX')
IF(SEX EQ 'MALE' OR SEX EQ 'FEMALE'), JUMP LAB2
PRINT 'SEX MUST BE MALE OR FEMALE'
JUMP LABEL1
LAB2:---
```

The following rules apply to use of the JUMP command for program branching control in a retrieval:

- a branch can always be made anywhere within a retrieval block.
- a branch can be made from within a block to outside of the block.
- a branch cannot occur from outside a block directly into the block but only to the block header (e.g. PROCESS REC, FOR EACH REC, SUMMARY, etc.)

NREAD ('prompt text')

A function which issues the specified prompt to the terminal along with a '?'. A numerical response is then read from the terminal in free format and placed into the command from which the function was executed.

```
10 COMPUTE VALUE = NREAD ('WHAT IS THE FIRST VALUE')
RECORD IS { rt# } [,(sort id1, sort id2,...)]
          { name }
```

Identifies a specific record to be created or retrieved within the present case. If a record is to be created, RETRIEVAL UPDATE is required.

```
RETRIEVAL UPDATE
CASE IS 127
RECORD IS 3,(1976,OCT)
COMPUTE SALES = 144.6;PROFIT=9.5;UNITS=1552
FINISH
```

In this example, a new record type 3 record is created in the case with id 127 and with sort variable values 1976 and OCT. This new record will contain the indicated values for SALES, PROFIT and UNITS.

SREAD ('prompt text')

A function which issues the specified prompt to the terminal followed by a '?'. A single character string response is then moved from the terminal into the command in which the function call is located.

SIR - SUPPLIED PROCEDURES

10 LABEL1 : COMPUTE NAME=SREAD('ENTER LAST NAME')

SUPPLIED PROCEDURES=====

The SIR procedures are an important part of the SIR retrieval capability. They enable the user to:

- obtain simple descriptive statistics on the summary variables (FREQUENCIES, PLOT, CONDESCRIPTIVE),
- interface directly to powerful statistical packages (SPSS SAVE FILE, BMDP SAVE FILE),
- create a new SIR file from the summary records (SIR SAVE FILE),
- create a printer or card image file from the summary records (WRITE RECORDS),
- produce a complex, hierarchical report from the summary records (REPORT).

The procedure commands appear within the body of the user's retrieval command set between the SUMMARY and AFTER RETRIEVAL blocks. If the SUMMARY command has an EXCLUDE or INCLUDE option specified, only the remaining summary variables are available to procedures unless otherwise specified by a VARIABLE clause in the procedure.

```
BMDP SAVE FILE      FILENAME = filename/
                    VARIABLES = { ALL
                                variable list } /
                    SORT = variable list /
                    SAMPLE = proportion [,seed]/
                    BOOLEAN = (logical expression) /
```

Creates a BMDP save file of the summary records.

```
BMDP SAVE FILE      FILENAME = BMDPRN/SAMPLE = .1
CONDESCRIPTIVE      VARIABLES = { ALL
                                variable list } /
                    WEIGHT = variable name/
                    SAMPLE = fraction [,seed]/
                    TITLE = 'text'/
                    BOOLEAN = (logical expression)/
                    STATISTICS = { ALL
                                   keyword list } /
```

Provides a listing of any or all of the following statistics:

VAR...variance	SKEW...skewness
MAX...maximum value	KURT...kurtosis
MIN...minimum value	CV.....coef of variability
MEAN...the mean value	CI.....95% confidence interval
STDE...standard error of	about the mean
the mean	WCOUNT..weighted count of non-
STDV...standard deviation	missing values.

```
CONDESCRIPTIVE      VARIABLES = INCOME TO SOCSCR/
                    BOOLEAN = (GROUP EQ 1)
                    STATISTICS = MAX, MIN, MEAN, STDV
FREQUENCIES         INTEGER = { ALL
                                variable list } (low,high)
                    [variable list(low,high)...]/
or GENERAL = { ALL
              variable list } (# of cells)[...]/
or CONTINUOUS = { ALL
                 variable list } (# of cells,
                                low,high)
                 [variable list...]/
or INTERVALS = { ALL
                variable list } (interval values)
                [...] /
WEIGHT = variable name/
```

```

SAMPLE = proportion/[ ,seed]/
BOOLEAN = (logical expression)/
STATISTICS = { ALL
              { keywords } } /
TITLE = 'text' /
HISTOGRAM /

```

Provides frequency tables and optionally histograms on desired variables. The statistics option list available is that of the CONDESCRIPTIVE procedure with the addition of:

```

MED ..... median
MODE ..... mode

FREQUENCIES GENERAL = SEX(2), SOCCCLASS(6)/
HISTOGRAM
FREQUENCIES CONTINUOUS = BLOODPRS (8,60,200)/
BOOLEAN = (ADULT EQ 1)

PLOT

SCATTERGRAM = dependent variable list (low,high) WITH
or           independent variable (low,high) BY control
              variable (low, high)/
LINEPLOT     = dependent variable list (low, high) WITH
              independent variable (low,high) BY control
              variable (low, high)/
TITLE        = 'plot title'/
TITLEX       = 'x-axis title'/
TITLEY       = 'y-axis title'/
XDIV         = number of x-axis divisions/
YDIV         = number of y-axis divisions/
SYMBOLS      = symbol list/
SAMPLE       = proportion[ ,seed]/
BOOLEAN      = (logical expression)/

STATISTICS   = ALL or NONE/
DEVICE       = PRINTER (number of columns, number of lines
                  lines per page)/

```

Provides two way scattergram or line plots of two variables with the option of stratifying the plots on a third variable. A different symbol will be used for each stratum. Statistics printed when requested include the following results of a simple linear regression of the y-axis variable on the x-axis variable:

- correlation coefficient, r
- r-squared
- significance of r
- standard error of the estimate (std dev of residuals)
- intercept, its std error and significance
- slope, its std. error and significance

```

PLOT SCATTERGRAM = AGE(20,80) WITH SYSBP(60,180)
      BY SEX (1,2)/
TITLE = 'SYSTOLIC BP WITH AGE BY SEX'/
TITLEX = 'PATIENTS AGE'/TITLEY = 'SYSTOLIC BP'
SYMBOLS = M,F/
STATISTICS = ALL

```

```

SIR SAVE FILE FILENAME = filename/
VARIABLES = { ALL
             { var list } } /
SAMPLE = proportion [ ,seed]/
BOOLEAN = (logical expression)/
RECTYPE = number [ ,name]/

```

Creates a new SIR file from the summary records. To allow changing information in the new SIR file from that found in the source SIR file code book file, the following additional keywords are available to indicate desired characteristics of the new SIR FILE.

SIR - REPORT GENERATOR

NOFCASES = number of expected cases/
 RECSCASE = estimated average records per case/
 CASEID = name of case id variable
 MAXTYPES = maximum number of record types/
 MAXKEYSZ = maximum # of bits in the key/
 MAXRECS = maximum records of any type/
 SORTIDS = defines a list of sort id variables
 when there are to be two or more
 records per case.
 INTEGER = variable list (min, max, scale),
 [variable list (min, max, scale)...]
 defines the list of variable to be
 stored as integers in the new SIR
 file. All variables not named in
 this command will be stored as real
 numbers.

SPSS SAVE FILE FILENAME = filename/
 VARIABLES = variable list /
 SUBFILES = varname (value1) name1(value2)
 name2.../
 SORT = variable [(D)][,variable...]
 WEIGHT = weight variable /
 SAMPLE = fraction [,seed]
 BOOLEAN = (logical expression)/

Creates an SPSS save file from the summary records.

SPSS SAVE FILE FILENAME = SPSSRUN/SUBFILES=SEX (1)MALE(2)FEMALE

WRITE RECORDS FILENAME = name/
 VARIABLES = variable list/
 FORMAT = (format specification)/
 SAMPLE = proportion[,seed]
 BOOLEAN = (logical expression)/
 SORT = variable [(D)] [,variable [(D)],...]/
 MISSCHAR = character

Writes the summary records to the named file. The format specifications may include:

Fw.d Floating point field, w columns wide, d columns to the
 right of the decimal point. Right justified.
 Iw Integer field, w columns wide. Right justified.
 Aw Character string field, w columns wide. Left justified.
 nX Skip n columns on the output record.
 Tn Tab to column n before printing the next value. Tabbing
 may only be done to the right.

Text in the FORMAT specification may be enclosed in quotes or be
 specified by a hollerith designation (eg 4HDATA). A slash indi-
 cates the start of a new line. A repeat multiplier may be used
 and parenthetical expressions may be nested to 10 deep. By
 default, when a variable has a missing value, the specified field
 is filled with blanks. The MISSCHAR specification provides a
 user specified alternative character to be used in filling
 missing value fields.

----- REPORT GENERATOR=====

The SIR report generator provides the user of SIR with the tools need-
 ed to produce complex, hierarchical reports. It is a SIR procedure
 and works in a manner similar to the other SIR procedures, i.e. it
 operates on one summary record at a time, in the order they are pro-
 duced by the retrieval unless a SORT clause is specified. The REPORT
 procedure should be used when the report-generating facilities built
 into the SIR retrieval are inadequate. Consequently, the report
 procedure will be needed whenever the report requires sorting,

multiple levels of headings, sub total and grand total computatic output, and the ability to print different kinds of output for different records.

The only data that may be used in a report are the summary record variables and the variables defined within the report itself using COMPUTE and IF statements. The user may not refer to CIR variables or any variables in the various records of the SIR data base directly from within a REPORT procedure. The user's retrieval program, which creates the summary records available to the report, should be designed so that any CIR variables or record variables required in the report are transferred to the summary record.

The special variables PAGE, DATE and TIME are always available to the report procedure.

In a report, we want certain actions to take place when the values of specific variables or logical conditions change. We call these changes breakpoints. Usually the action we desire at breakpoints is printing subtotals and grand totals. The structure of the report is determined by the ordering of the records from which the report is to be created on the breakpoint variables. The user decides in advance which variables constitute the breakpoint variables and orders the summary records appropriately on these variables by use of the SORT clause in the REPORT command.

```
REPORT                SORT = variable[(D)][,variable[(D)]...]/
                     BOOLEAN = (logical expression)/
                     FILENAME = local file name/
                     SAMPLE = proportion/[,seed]/
                     PAGESIZE = number of lines per page[,columns
                               per line]/
```

The FILENAME clause is required, all other clauses are optional. The default for the PAGESIZE command is 60, 136 respectively. The number of columns per line must lie within the range 72-136 and the number of lines per page cannot exceed 66. The variable appearing in the BOOLEAN statement must be a summary variable only. The SIR REPORT procedure command is followed by a command set which produces the desired report. In addition to the following special report commands, HEADING, FOOTING, WRITE, WRITE IF, COMPUTE, BLANK LINE(S), PAGE EJECT and IF commands may be used as required.

AFTER REPORT

Begins a block of commands placed at the very end of the report procedure to be executed once after the report is complete.

```
AFTER REPORT
WRITE 'THE REPORT IS COMPLETE...THERE WERE',COUNT(15),'CASES
      WRITTEN'
END
```

AT END

Begins a set of commands to be executed each time a higher level sort variable changes in value and before execution is allowed to return to the higher level. It is located after the LEVELS heading and computation commands. It provides for printing subtotals at breakpoints. If there are several AT END sets in various levels of a branch, they are executed in reverse order. Summary variables should not be referenced in AT END sets directly as they will contain the new value which generated the breakpoint. If it is necessary to print the value of a summary record variable as it exists prior to the breakpoint, it must be first saved in a new variable by a COMPUTE statement prior to the AT END set. The saved value can then be referenced in the AT END set.

```
LEVEL                2,FAMMEM
COMPUTE              MEMTOT=0
AT END
WRITE                40T,'TOTAL',48T,COST('$9999.99')
```

SIR - REPORT GENERATOR

BEFORE REPORT

This block of commands may be used, immediately after the REPORT command to provide an introductory section to a report. This block is executed once at the very beginning of the report before any summary records are read. No variables from the summary record can be referenced in this block as they are not yet available to the REPORT procedure. BEFORE REPORT is designed for printing a cover page for the report and for initializing global variables for the report.

DETAIL

Begins a set of commands for forming the body of the report. Each time a particular branch containing a DETAIL set is executed the commands in the DETAIL set are executed. A DETAIL block is usually located at the innermost level of each branch of a report although it may be located within any level. The DETAIL section will usually contain commands for printing individual lines of reports and for incrementing subtotals and grand totals.

```
DETAIL
COMPUTE          VDATE=JULC(VISTDATE)
WRITE           6T,VDATE(A12),31T,PROCDR(14),50T,COST(F6.2)
COMPUTE          GRTOT=GRTOT+COST;MEMTOT=MEMTOT+COST
```

```
END
END LEVEL
END REPORT
```

To improve readability of the user's REPORT procedure, an END command may appear at the end of BEFORE REPORT, AFTER REPORT blocks or at the end of LEVEL sections within the body of the report. The END cannot be used to separate the three sections of the LEVEL-headings, DETAIL and AT END. Optionally, the statement END LEVEL may be used to end a LEVEL section and END REPORT may be used to terminate the report set.

```
EXIT
EXIT IF          (logical expression)
```

Terminates the report. Can be used in the ON ERROR level to terminate the report in case of error.

```
JUMP TO LEVEL   level identification
```

Directs execution from one branch path into another at a higher level where necessary report commands have already been defined. Allows re-joining separate paths to avoid multiple sets of duplicate commands.

```
LEVEL           level number           ,breakpoint variable name
                                           (breakpoint logical con-
                                           dition)
```

Identifies the levels in the nesting structure of the report. Nesting levels are indicated hierarchically by the level number. Levels are identified numerically 1 highest, 2 next...,k, lowest or innermost. Multiple branches at the same level are identified by decimal numbers m.n(eg. 2.1,2.3,2.4,etc.). Multiple branch modes in different branches which are identical may be labeled with up to 3 characters for identification(2.1A,2.1B. 2.1ABC.etc.).

The breakpoint variable name must be the name of a summary record variable. A change in the value of this variable triggers a breakpoint action to occur at the given level. All breakpoints nested within the given level also take place.

The breakpoint logical condition is a standard SIR logical expression. A change in the value of this logical expression from false to true, or vice versa triggers a breakpoint at the given level and all levels nested within it. When a logical condition is used on a LEVEL command, the level number must be of the form

m.n to indicate one of multiple possible branches at the same level within a report.

Any computation or output commands which appear immediately after the LEVEL command and which are not contained within an AT END or a DETAIL command set will be executed only when a breakpoint occurs at that level or at a higher level.

```
LEVEL          1,FAMID
LEVEL          2,FAMEM
LEVEL          3.1,(FAMEM EQ 1)
LEVEL          3.2,(FAMEM EQ 2)
LEVEL          3.2A,(STATUS LT 4)
```

ON ERROR level number

The ON ERROR command provides a special LEVEL command which can be used in multiple branch reports where the LEVELS commands present in the report do not cover all possible conditions that could occur. If a condition occurs that is not covered by at least one of the branches, SIR would normally abort the report execution. The ON ERROR command will be executed instead if it is present. If a level number is included, it must be of the form m.n and applies only to errors occurring at level m. If no level number is included, it is a global error recovery block and applies to errors encountered at any level not covered by an ON ERROR block at a specified level. If used without a level number, this set can appear anywhere in the body of the report.

```
ON ERROR      3.2
WRITE        'ERROR AT LEVEL 3...CASE',FAMID(16)
EXIT
END
```

ON PAGE level number

Tells the system to which level to return for printing headings when a page eject occurs. When a PAGE EJECT command is encountered or an automatic page eject condition occurs, and an ON PAGE command has been included in the report, the following actions take place:

- The page footing is printed if a FOOTING command was included.
- Page eject is performed.
- The HEADING command, if any, is executed.
- Starting at the level specified in the PAGE command, the current branch of the report is followed from the outermost level and all WRITE, WRITE IF and BLANK LINES statements which are not in DETAIL and AT END sections are executed. All other commands are ignored.
- When the statement that caused the page eject is reached, the report processing resumes in the normal manner.

If there is no ON PAGE command in the report, only the first three steps above will be taken before processing continues.

```
ON PAGE      3
```

SIR UTILITIES=====

SIR utilities provide capabilities for maintaining, archiving and modifying the SIR data base. To run a utility, appropriate PASSWORD, SECURITY (with data base administrator passwords) and GET FILE commands are required.

SIR - UTILITIES

TAPE STORE

FILENAME = local filename/NOREWIND/
NEWFN = newfile name/
NEWPW = new password/

This utility moves the entire data base (ie. code book, detail, procedures and data file) to the named local file. This file will normally be a magnetic tape although there is no restriction on the physical location of the resulting file. If the NOREWIND option is specified, the file written will not be re-wound at the end. This utility does not purge the existing SIR data base. This utility is also used to unlock a locked data base.

PURGE SIR FILE

This utility deletes the SIR data base from the disk. It will typically be used after the data base has been stored on magnetic tape by use of the TAPE STORE utility.

SIR MERGE

FILENAME = source SIR filename/
PASSWORD = source SIR password/
SECURITY = source SIR read password/
RECTYPE = source #, host #, host name/
IDTRUNC = YES or
 NO /
UPDATE = ADD or
 REPLACE /
BOOLEAN = (logical expression)/
RENAME = (old var name, new name), [(...)]/

Merges a single record type of a source SIR file into a host SIR data base which may contain multiple record types.

The first four clauses FILENAME, PASSWORD, SECURITY and RECTYPE are required and in the order shown. The SECURITY commands must provide level 30 password access to the source SIR data base. The RECTYPE clause indicates the present record type number of the source record type and the number and name of the record type that it will be in the host SIR data base. IDTRUNC is used to truncate string case id's which may be longer than those allowed in the host SIR data base. UPDATE is used when the source record type is being merged with an existing record type in the host. If UPDATE is not specified, both adding new records in the host and overwriting old records is allowed.

If the source record type is merged into an existing record type in the host, the following restrictions must be met:

- The source and host record types must have identical sort id lists.
- Each variable in the source record must have the same type and length as the host. Missing values defined in the host are used in all cases.
- VAR RANGES and VALID VALUES are ignored.
- Only the VAR LABELS and VALUE LABELS of the host remain as the labels in the host after the merge.

The RENAME clause allows renaming source variables to avoid duplicate names with those already in other record types within the host SIR data base.

The BOOLEAN can reference variables in the source SIR data base only.

SIR SUBSET

FILENAME = filename/
RECTYPES = rt# or name, (logical exp.)[,rt# or
 name, rt# or name, (log.exp.),...]/
BOOLEAN = (logical expression)/
SAMPLE = fraction [,seed]/

```

or COUNT = total[increment [,start]]/
or LIST  = list of case identifiers/

```

Creates a new SIR file which is a subset of the original SIR file. Specific record types or parts thereof can be extracted. The BOOLEAN can apply to the CIR record only. The RECTYPES clause specifies the list of record types that are to be included in the new subset SIR data base and any logical filters that are to be applied in deciding which records from the specified record types are to be included in the subset. If a logical expression is included for a record type, it can only reference variables in that record type or the CIR. The COUNT, LIST and SAMPLE clauses are as defined for the FOR EACH CASE command.

```

SIR FILE DUMP      FILENAME = filename/[NOREWIND/]
                  SEQUENCE = begin column, end column /
                  RECTYPES = rt# or name,(logical expression)
                        [,rt# or name,rt# or name (logical
                        expression),...]/
                  BOOLEAN = (logical expression)/
                  SAMPLE  = fraction [,seed]/
or COUNT          = total, increment, start/
or LIST           = list of case identifiers /
                  COMPVAR = NO
                        YES /

```

Produces a dump of the SIR data file to a card image file which is as closely as possible a reproduction of the data records originally entered. An abbreviated code book schema definition command set is first written to the specified file followed by the requested data records (for a complete schema command set, use the WRITE SCHEMA command). Computed variables created at data entry or by SIR MERGE will conform to the following rules when written to cards:

- Computed variables, if numerical, will be output on subsequent cards at the end of the record type to which they refer. They will be written in F10.n format in columns 1-72. E conversion will be used as necessary if the number will not fit using F1.n.
- Computed variables of character type will be output on subsequent cards following the numerical computed values.
- The specification COMPVAR = NO suppresses the output of computed variables and should be used if only a reproduction of the original data is desired.

Note that any variables added to the SIR data base by an SIR MERGE are considered computed variables.

```

SIR FILE LIST      RECTYPES = { rt# or name,(log.exp)[,rt# or name
                        rt# or name,(log.exp),...]}
                  ALL
                  ORDER   = { ALPHA
                        VARNUM } /
                  BOOLEAN = (logical expression)/
                  COUNT   = total[increment[,start]]/
or LIST            = case id [,case id ,..., [case id
                        THRU case id ]] /
or SAMPLE          = fraction [,seed]/

```

Produces a sequential listing, record by record of the data values present within each selected record within the requested record types of the desired cases. The ORDER clause controls the order in which variables are printed within each record. ALPHA implies alphabetical ordering and VARNUM implies variable number ordering. VARNUM is the default.

SIR EXAMPLES=====

CODE BOOK SCHEMA DEFINITION EXAMPLE

SIR - EXAMPLES

This code book schema definition command set defines a two record type SIR data base made up of demographic and clinic visit records on patient cases.

```

RUN NAME      EXAMPLE DATA BASE DEFINITION RUN
C             ***SYSTEM INITIALIZATION COMMANDS
FILE NAME     EXAMPLE
PASSWORD      AAAA
READ SECURITY  (30)MASTERR,(15)USER1R,(5)USER2R
WRITE SECURITY (39)MASTERW,(15)USER1W,(5)USER2W
C             ***CASE DEFINITION COMMANDS***
TASK NAME     CASE STRUCTURE DEFINITION
N OF CASES    500
RECS PER CASE 150
MAX REC TYPES 10
CASE ID       PATNTNUM
RECTYPE COLS  5,6
SEQUENCE COL  80
DOCUMENT      THIS EXAMPLE ILLUSTRATES COMMANDS FOR THE DEFINING
              OF A STUDY DATA SET CONSISTING OF TWO RECORD TYPES.
COMMON LIST   PATNTNUM,BRTHDATE,SEX,VISTDATE
C             ***RECORD TYPE 1 DEFINITION SET***
TASK NAME     RECORD TYPE 1 -DEMOGRAPHIC DATA-
RECORD SCHEMA 1,DEMOGRPH
DOCUMENT      DATA FOR THIS RECORD IS COLLECTED AT THE PATIENTS
              FIRST CONTACT WITH THE CLINIC.
MAX REC COUNT 1
VARIABLE LIST PATNTNUM,NAME,ADDRESS,BRTHDATE,DOCTOR,AGE,SEX,
              OCCUP,EDUCTN,INCOME
              (I4,1X,A25,5X,A40,/,5X,A6,2I2,3I1,I6)
INPUT FORMAT  BRTHDATE(MMDDYY)
DATE VAR LIST
VAR LABELS    PATNTNUM ,THE PATIENTS ID NUMBER/
              NAME ,THE PATIENTS NAME/
              INCOME ,FAMILY INCOME
VALUE LABELS  OCCUP (1)BLUE COLLAR (2)HOUSEWIFE (3)WHITE COLLAR
              (4)OTHER
VALID VALUES SEX(1,2)
MISSING VALUES NAME TO INCOME (BLANK)
VAR SECURITY   NAME, ADDRESS(30,30)
C             ***RECORD TYPE 2 - CLINIC VISIT DATA***
TASK NAME     RECORD TYPE 2 - PATIENT CONTACT
RECORD SCHEMA 2,CONTACT
DOCUMENT      THIS RECORD IS COLLECTED EACH TIME THE PATIENT MAKES
              A HEALTH CONTACT WITH THE CLINIC.
MAX REC COUNT 300
SORT RECORDS  VISTDATE,VISTTIME
DATA LIST     FIXED(1)/1
              PATNTNUM 1-4 (I),VISTDATE 5-10 (A)
              VISTTIME 11-14 (A),WEIGHT 15-19 (I)
              HEIGHT 20-21 (I)
DATE VAR LIST VISTDATE(MMDDYY)
TIME VAR LIST VISTTIME(HHMM)
SCALED VAR LIST WEIGHT (-2)
VAR RANGES    WEIGHT(30.00,160.00)/HEIGHT(20,80)
REJECT REC IF (VISTDATE LT BRTHDATE)
TASK NAME     CODE BOOK LISTING REQUESTS
LIST SCHEMA   STRUCTURE, LABELS,DETAILED
    
```

BATCH RETRIEVAL

This retrieval command set produces an SPSS SAVE FILE made up of clinic contact data retrieved from the above SIR data base on patients over 21 years of age at entry who are overweight.

```

RUN NAME      EXAMPLE SUMMARY RECORD RETRIEVAL
GET FILE      EXAMPLE
PASSWORD      AAAA
SECURITY      USER1R, USER1W
RETRIEVAL     PRESET
PRESET        RECCNT(0)
C             ***READ THE DEMOGRAPHIC RECORD***
PROCESS REC   1
C             ***MOVE DESIRED VARIABLES***
MOVE VAR LIST BRTHDATE TO INCOME
    
```

```

END
C          ***DECIDE IF CASE IS DESIRED
REJECT CASE IF (AGE LE 21)
C          ***READ CONTACT RECORD***
FOR EACH REC 2
C          ***IF OVERWEIGHT,CREATE SUMMARY REC***
SELECT REC IF ((SEX EQ 1 AND WEIGHT GT 85) OR
              (SEX EQ 2 AND WEIGHT GT 68))
COMPUTE RECCNT=RECCNT+1
MOVE VAR LIST VISTDATE TO INCOME
END
SUMMARY EXCLUDE RECCNT
SPSS SAVE FILE FILENAME=OVERWT
AFTER RETRIEVAL
WRITE '***SUMMARY COMPLETE',/,5X,RECCNT(I5),'RECORDS
      WRITTEN'
FINISH

```

SIR REPORT EXAMPLE

Consider a data file made up of family clinic records. A report is desired which lists procedures done by Dr. Jones for his patient families, listing each procedure, the date it was done and its cost. The report is to be organized by family broken down by father, mother and children. A sample from the desired report is as follows:

ANNUAL PATIENT COSTS - DOCTOR JONES	1/15/78	PAGE NO...1
FAMILY NAME ANDREWS		
ADDRESS 1632 S 28TH ST SEATTLE		
HUSBANDS NAME JOHN		
DATE	PROCEDURES	COST
FEB 13,1977	3906	16.00
DEC 1,1977	1312	6.20
	TOTAL	\$47.60
WIFES NAME SUE		
DATE	PROCEDURES	COST
MAR 5,1977	4218	25.96
OCT 14,1977	3762	125.76
	TOTAL	\$151.72

In the data file, the CASE ID variable identifies individual families and its name is FAMID. Within the families, individual members are identified by a sequential number (i.e., father = 01, mother = 02, children = 11, 12, etc.) stored in the variable FAMMEM. Finally, each office visit is recorded by date in the variable VISTDATE, which is stored as a julian integer. We must first order the records by families (FAMID), then by family members (FAMEM) within cases and finally select only those visits for the year 1977 and for which JONES was the doctor. The following SIR command set is sufficient to produce the desired report:

```

RETRIEVAL
FOR EACH REC VISIT
SELECT REC IF (DOCTOR EQ 3 AND (VISDATE GT JULN
                              (01,01,77) AND LT JULN (12,31,77)))
MOVE VAR LIST VISTDATE, PROCDR, COST
PROCESS REC DEMOGRPH
MOVE VAR LIST FAMID, FAMMEM, LNAME, FNAME, STREET, CITY
REPORT SORT = FAMID,FAMMEM,VISTDATE/
        FILENAME = REPORT1
BEFORE REPORT HEADING 'ANNUAL PATIENT COSTS - DOCTOR JONES'
                     40T, DATE, 'PAGE NO..',PAGE
LEVEL 1,FAMID
WRITE 'FAMILY NAME',20T,LNAME(A20)/'ADDRESS',20T,
      STREET, CITY
COMPUTE GRTOT = 0;HEADING = 'CHILDREN'
AT END

```

SIR -EXAMPLES

```

WRITE          20T, 'GRAND TOTAL', 35T, GRTOT('$9999.99')
LEVEL         2,FAMMEM
COMPUTE       MEMTOT = 0
AT END
WRITE        40T, 'TOTAL', 48T, MEMTOT ('$9999.99')
LEVEL       3.1(FAMMEM EQ 1)
WRITE      5T, 'HUSBANDS NAME', 22T, FNAME(A10)
WRITE      8T, 'DATE', 26T, 'PROCEDURES', 52T, 'COST'
LEVEL      4
ON PAGE     2
DETAIL
COMPUTE     VDATE = JULC (VISTDATE)
WRITE      6T, VDATE(A12), 31T, PROCDR(I4), 50T, COST(F6.2)
COMPUTE     GRTOT = GRTOT + COST; MEMTOT = MEMTOT + COST
LEVEL      3.2, (FAMMEM EQ 2)
WRITE      5T, 'WIFES NAME', 22T, FNAME(A10)
JUMP TO LEVEL 4
LEVEL      3.3, (FAMMEM GT 2)
IF          (HEADING EQ 'CHILDREN') WRITE 5T, HEADING(A8)
COMPUTE     HEADING = ' '
WRITE      22T, FNAME(A10)
WRITE      8T, 'DATE', 26T, 'PROCEDURES', 51T, 'COST'
JUMP TO LEVEL 4
ON ERROR    3.4
WRITE      'ERROR IN CASE', FAMID(I6),/
           'FAMILY MEMBER', FAMMEM(I2),
           'VISIT DATE', JULC(VISTDATE)

EXIT
END REPORT
    
```

INTERACTIVE RETRIEVAL

Using the same SIR data file from which the REPORT example above was done, suppose we wish to be able to interactively request a list of all visits by date, procedure done and cost for any patient in the data base. We wish the program interaction and report to look like the following:

RUN PROCLIST

THIS PROGRAM LISTS A REPORT ON ALL
PROCEDURES DONE-TO-DATE ON ANY
PATIENT IDENTIFIED BY FAMILY ID
AND MEMBER ID

WHAT IS THE FAMILY ID? 178463
INDIVIDUAL ID? 01

REPORT ON JOHN SMITH # 178463 JAN. 1,1978
ADDRESS 1832 14TH S.W.
 SEATTLE WN
 93841
PHONE (206)-842-1853

DATE	PROCEDURE	COST
FEB 13,1977	3906	16.00
OCT 14,1977	4219	9.00

TOTAL		25.00

CONTINUE? YES
WHAT IS THE FAMILY ID?

The following SIR command set will provide the above interactive report when save as proclist in the SIR procedure file.

```

10 RETRIEVAL
20 WRITE        1X,/, "THIS PROGRAM LISTS A REPORT ON ALL",/,
30                "PROCEDURES DONE-TO-DATE ON ANY",/,
    
```



```

40          "PATIENT IDENTIFIED BY FAMILY ID",/,
50          "AND MEMBER ID",/
60 L1 : COMPUTE IDENT = NREAD ("WHAT IS THE FAMILY ID")
70          ;INDID = NREAD ("INDIVIDUAL ID")
80          ;COSTT = 0.0
90 CASE IS IDENT
100 PROCESS REC DEMOGRAPH, WITH (INDID)
110 WRITE      /,"REPORT ON",2X,FNAME, LNAME,
120          2X,"#",1X,IDENT (I7),40T,DATE
130 WRITE      5T,"ADDRESS",15T,STREET,/,
140          15T,CITY,STATE,/,15T,ZIP,/,
150          5T,"PHONE",15T,PHONE,/
160 WRITE      4T,"DATE",20T,"PROCEDURE",
170          35T,"COST",/
175 PROCESS REC VISIT,WITH(INDID)
180 COMPUTE    PDATE = JULC(VISDATE);COSTT = COSTT + COST
190 WRITE      PDATE(A12),22T,PROCDR(I4),33T,
200          COST(F6.6)
210 END
220 WRITE      30T, '-----',/,25T,'TOTAL',
230          33T,COSTT ('$999.99')
240 COMPUTE    TEST = SREAD ('CONTINUE')
250 IF        (TEST EQ 'YES' OR 'Y') JUMP L1
-----

```

SIR EXECUTION OPTIONS=====

The SIR command has the general form:

SIR, p1,p2,...,pN

where the p's stand for parameters listed below. These parameters may appear in any order within the command.

Some of the parameters have default and alternate default values. When the parameter does not appear in the list, the default value is used. When it appears with no value specified, the alternative default is used. When it appears in the form p1=value, the value given is used for the parameter.

Parameter	Type	Default	Alternate
I = input file	Batch	INPUT	COMPILE
L = output file	Batch	OUTPUT	OUTPUT
IA	Interactive	-	-
FN = SIR file name	Interactive	-	-
PW = password	Interactive	-	-
RS = read security	Interactive	-	-
WS = write security	Interactive	-	-
PS = page length (lines)	Interactive	60	24
WD = page width (cols.)	Interactive	136	80
EP = procedure file name	Interactive	-	-

In the batch mode, the command
SIR.

is the same as:

SIR, I=INPUT, L=OUTPUT.

In the interactive EDITOR mode, the command:

SIR, IA.

is the same as:

SIR, IA, PS=60, WD=136.

The IA parameter must be used to obtain the SIR Editor subsystem. It is suggested for terminal usage that the command:

SIR, IA, PS, WD.

be used to provide the alternate default values of 24 lines/ page and 80 columns/line, appropriate for many CRT terminals.

When using SIR interactively, any required parameter (eg. PW, RS, WS, etc...) not appearing in the SIR execution command will be requested via an interactive prompt. Consequently, the more parameters the user includes in his execution command statement,

SIR - EXECUTION OPTIONS

the fewer introductory prompts he need contend with to get going in SIR.

With this in mind, the use of an executable catalogue procedure can reduce execution of SIR to simply typing a single user chosen word after completing logon procedures. For example, suppose the catalogue procedure looks like the following:

```
GRAB(SIR)
SIR, IA, FN = FSTUDY, PS=10,WD=72,SP.AR, SP.
PW = JON, RS=DBAR, WS=DBAW
```

Provided the user has stored these lines as a catalogued procedure file called SIREX, he may bring up SIR on the study FSTUDY with the single command:

```
-SIREX
GOOD MORNING
WELCOME TO SIR 1.1
SIR EDITOR READY
>
```

During SIR interactive runs, SIR prints the following "pause message" at the end of each page:

```
ENTER C/H/N/S/BLANK CR >
```

C response tells SIR to continue the listing without further pausing. It is equivalent to setting the page size to zero for this listing. Blank followed by carriage return will continue printing the next page after which it will again pause. The "N" and "S" specifications are useful when the output from a SIR run (as opposed to an Editor listing) is being printed at the terminal. Both "N" and "S" responses terminate further output at the terminal and return control to the SIR Editor. However, "N" allows the SIR run to be completed (without further listing at the terminal), while "S" terminates the SIR run immediately with no further processing occurring whatever before control is returned to the editor.

On most computer systems, the execution of SIR may be interrupted from its present task and returned to the SIR Editor mode by simply typing the BREAK key on the keyboard. When the user types BREAK, the messages:

```
SIR JOB REPRIEVE
SIR EDITOR READY
>
```

are typed on the terminal. All execution has been terminated and the user may proceed with another task. This capability should be used with caution and as a 'last resort' as it unconditionally terminates whatever is presently being done. It is useful, however, when the user finds he has begun something he does not wish to continue with. The user should check with his local site about alternatives to the BREAK key that might be in affect on his system.

When additional files are to be used by SIR during a run (eg. if data is to be read), these files must be attached as local files under the name referenced for them in the SIR run before executing SIR (or the \$ command may be used, in SIR, to attach them).

The four SIR data base files (code book, detail, procedure and data) are saved automatically by SIR as permanent files when they are created during the SIR creation run. Any other files created by the SIR run (eg., an SPSS save file or an output list file) will exist after SIR has been terminated as local files under the name given to them in the SIR run. They must be properly disposed of (listed, saved, etc.) by the user prior to logoff or they will disappear.

SPSS GENERAL RULES =====

The contents of the control field must begin in column 1. The control field is columns 1-15. The specification field is columns 16-80.

File names and variable names may have a maximum length of 8 characters, the first of which must be alphabetic. They must be composed of letters and/or numbers. No special characters or embedded blanks may be used.

The following words may not be used as variable names.

ABS	EQ	MOD10	SIN
ACCUM	GT	NORMAL	SQRT
ALL	JULIAN	NE	SUBFILE
AND	LAG	NOT	TO
ATAN	LE	OR	TRUNC
BY	LG10	POISSON	UNIFORM
CASWGT	LN	RND	WITH
COS	LT	SEQNUM	EXP

If sequence numbers are present in columns 73-80 of the control statements, NUMBERED YES must be specified.

The maximum number of variables per system file is 1000. The maximum number variables per archive file is 5000.

The maximum number of subfiles per file is 100.

The MISSING VALUES statement can specify a maximum of 3 values per variable.

The ASSIGN MISSING statement can specify only one missing value per variable.

SUBFILE, SEQNUM, CASWGT values cannot be altered through COMPUTE, IF and RECODE statements or their starred counterparts.

Valid formats on the INPUT FORMAT statement are: A, F, P, S, T, V, X, /

The N of CASES statement must specify:

the exact number of cases for CARD input. If the exact number is not known use UNKNOWN option with an END OF RECORD card inserted at the end of the card input data.

Quotation marks must be used with alphanumeric values in:

- the control cards.
- the data if the FREEFIELD form of the INPUT FORMAT control card is used.

The maximum field width of an alphabetic variable is A10. However, only the first 8 characters of an alphabetic value will be printed.

SPSS, Statistical Package for the Social Sciences, Second Edition, Nie, Hull, Jenkins, Steinbrenner and Bent, McGraw-Hill, 1975.

SPSS - CONTROL COMMANDS

SPSS CONTROL COMMANDS =====

CONTROL WORDS	PARAMETERS
ADD CASES	n
ADD DATA LIST	FIXED(nrecs)/recnum varlist cols (format) (vartype) varlist... BINARY(nrecs)/recnum varlist varpos varlist...
ADD SUBFILES	subfile name(n1)...
ADD VARIABLES	varlist
ALLOCATE	TRANSPACE=number of bytes
ASSIGN MISSING	varlist(value)
COMMENT	any text
COMPUTE	computed variable =
*COMPUTE	arithmetic expression
COUNT	varname=criterion varlist
*COUNT	(value list)
DATA LIST	{ FIXED[nrecs]/recnum varlist starting cols [ending cols (vartype)] varlist... BINARY(nrecs)/recnum varlist varpos varlist }
DOCUMENT	any text
DELETE SUBFILES	subfile name list
DELETE VARS	varlist
DO REPEAT	stand-in varname=varlist/
END REPEAT	any comment
ECOLOGY	ON OFF
EDIT	check syntax only
ERASE DOC	remove documentation
FILE NAME	(file name, file label)
FINISH	terminate SPSS run
GET ARCHIVE	FILE=f1, VARIABLES=varlist
GET FILE	FILE= f5, VARIABLES=ALL file name
IF	(logical expression) computed variable=
*IF	arithmetic expression
INPUT FORMAT	FIXED(format list) FREEFIELD BINARY (format list)
INPUT MEDIUM	CARD DISK TAPE OTHER
KEEP VARS	var list
LIST ARCHINFO	VARLIST COMPLETE
LIST CASES	CASES=n/VARIABLES=varlist
LIST ERRORS	to get a listing of SPSS error messages (30 printed pages)
LIST FILEINFO	COMPLETE VARLIST SORTVARS VARINFO LABELS SUBDIRECTORY DOCUMENTS
MERGE FILES	FILES=f1 ,VARIABLES=ALL
MISSING VALUES	FILE= f5, VARIABLES=ALL varlist(missing value list)/
N OF CASES	n
NUMBERED	UNKNOWN YES NO
OPTIONS	option number list
OSIRIS VARS	variable number list
PAGESIZE	number of lines
	NOEJECT

one or more

```

PRINTBACK { NO
           CONTROL
           FORMAT
           varlist(value)/...
           file number

PRINT FORMATS
RAW OUTPUT UNIT
READ INPUT DATA
READ MATRIX
RECODE } varlist(values=newval)
*RECODE } varlist (CONVERT)/...
REJECT IF (logical expression)
RENAME VARS varlist=namelist/
           varlist=namelist...

REORDER VARS varlist
RUN NAME run label of up to 64 char.
RUN SUBFILES { (subfile list) (...)...
              EACH
              ALL
              }

SAMPLE
SAVE ARCHIVE sampling factor
              [compressed] [file name,label/]
              FILE= f1, VARIABLE=vrlst/

.
.
.
SAVE FILE FILE= f5, VARIABLES=ALL
SEED [compressed] [file name,file label/]
     STANDARD
     a large odd integer
     'alphabetic string'
     (logical expression)

SELECT IF varlist { (A)
*SELECT IF { (D)
SORT CASES }

STATISTICS { statistics number list
            ALL
            subfile name1(n1)
            label
            varlist(value1)label 1.../...
            varname1, var label/...
            variable list
            variable name

SUBFILE LIST
TASK NAME
VALUE LABELS varlist(value1)label 1.../...
VAR LABELS varname1, var label/...
VARIABLE LIST variable list
WEIGHT } variable name
*WEIGHT }

WRITE CASES { (format list)varlist
              OR
              BINARY varlist
              (options 1 with WRITE CASES
              gives listwise deletion of
              missing data.)
              }
VARIABLES=varlist
any one or comb, of:
      FILENAME PRNTFMTS
      VARLIST VARLABS
      SUBFLIST VALLABS
      NCASES DOCUMENTS
      MISVALS ALL
      }

CHAR { BCD
      EBCD
      }
NUMBERED

```

SPSS STATISTICAL PROCEDURES =====

```

-----
AGGREGATE GROUPVARS= varlist/
          VARIABLES= varlist/
          { any one or comb. of
            SUM
            VALIDN
            MEAN
            SD
            MAX
            MIN
            SKEW
            KURT
            PCTGT (value)
            PCTLT (value)
            PCTBTN (low,high)
            }
          [RMISS=value]/
AGGSTATS=

```

SPSS - ANOVA

Program to group cases on the basis of one or more aggregating variables.

AGGREGATE GROUPVARS=CITYID/
 VARIABLES=FREQUOTE,NCONTACT,CAMPAIGN/
 AGGSTATS=VALIDN,MEAN,SD

- OPTIONS: 1. Includes missing data.
 2. Deletes missing data, listwise.
 3. Grouping variables included in aggregated file.
 4. Outputs a compositional aggregated file.
 5. No end-of-record after each case in output binary file.
 6. An end-of-file to be written after the data on the binary output file.
 7. Suppresses the AGGREGATE output data file.
 8. Suppresses the printed output.
 9. Causes AGGREGATE output to be written on a formatted BCD output file.

- STATISTICS:
 1. Group identification number, total number of cases and values of grouping variables.
 2. Statistics 1, plus exact contents of each case in aggregate file.
 3. Statistics 2 for the first 10 cases only.
-

ANOVA dependent varlist BY independent varlist(min,max)
 WITH covariate list/dependent varlist.../...

One-to-five-way analysis of variance and covariance for factorial designs assuming a fixed-effects model.

ANOVA INCOME BY EDUC,RACE,SEX(1,3) WITH AGE,FATHINC

OPTIONS 8
STATISTICS 1

- OPTIONS: 1. Includes missing data.
 2. Suppresses value labels.
 3. Ignores two-way and higher interactions among factors.
 4. Ignores three-way and higher interactions among factors.
 5. Ignores four-way and higher interactions among factors.
 6. Ignores five-way and higher interactions among factors.
 7. Processes covariates concurrently with main effects for nonmetric factors.
 8. Processes covariates after main effects for nonmetric factors.
 9. Indicates regression approach.
 10. Indicates heirarchical approach.

- STATISTICS:
 1. Multiple classification analysis. (available with option 9.)
 2. Unstandardized partial regression coefficients for the covariates.
-

BREAKDOWN VARIABLES=varlist(low,high)/
 TABLES=varlist BY varlist BY... (Integer mode)
 OR
 TABLES= varlist BY varlist BY... (General mode)
 OR
 VARIABLES= varlist(low,high)/
 CROSSBREAK=varlist BY varlist... (Crossbreak mode)

Summary statistics among subgroups of the cases in the data file.

BREAKDOWN VARIABLES=ACHTEST(LO,HI)
APTEST1 TO APTTEST5(0,4)
MOTIVE/TABLES=ACHTEST
BY APTTEST1 TO APTTEST5
BY MOTIVE

OPTIONS: 2
STATISTICS: 1

- OPTIONS: 1. Includes missing data.
2. Excludes missing data for dependent variables.
3. Suppresses all labels.
4. Prints in tree-diagram format.

(CROSSBREAK Facility)

- 1-3 Same as BREAKDOWN
4. Not used.
5. Deletes cell frequencies.
6. Deletes the sums.
7. Deletes the standard deviations.
8. Deletes value labels but prints variables labels.

- STATISTICS: 1. One way analysis of variance table.
2. Test of Linearity table.(1 also required)

CROSSBREAK Facility

- 1-2 Not used.
3-12 Same as CROSSTABS 1-10 statistics.

CANCORR VARIABLES=varlist/
RELATE=(n,F,t) varlist WITH varlist/
{ CANVAR=mdrp } / [RELATE=...]
or
CANVAR
Canonical Correlation analysis program.

CANCOR VARIABLES=ATTITUD1 TO ATTITUD7
BEHAV1 TO BEHAV6/RELATE=ATTITUD1
TO ATTITUD7 WITH BEHAV1 TO BEHAV6/
CANVAR=.30/

OPTIONS: 2,10
STATISTICS: 3,4

- OPTIONS: 1. Includes missing data.
2. Deletes missing data, pairwise.
3. Computes a covariance matrix instead of a correlation matrix.
4. Matrix input.
5. Computes matrix about the origin.
6. Continues computations if either variable set is linearly dependent.
7. Continues computations if a matrix is not positive definite. (with option 2)
8. Outputs a simple correlation matrix file.
9. Uses VARIABLE LIST with matrix input. (with option 4)
10. Calculates weighted estimates of missing data in computing canonical scores. (with option 2)
11-17. Not used.
18. Reads in mean and s.d. with correlation matrix
19. Output canonical and variate scores file.
20. Substitutes means for missing data when correlation or covariance matrix is being computed

- STATISTICS: 1. Means and standard deviations.
2. Correlation matrix.
3. Correlation matrix only if non-computable coefficients are encountered.

4. Canonical variate scores.

```

CONDESCRIPTIVE  { varlist
                  }
                  ALL

```

Univariate summary statistics on continuous variables.

```

CONDESCRIPTIVE  INCOME TO SESINDEX,SCALE1,
                 SCALE3,ORGMEMBS,FAMSIZE

```

```

OPTIONS:        1,3
STATISTICS:     ALL

```

```

OPTIONS: 1. Includes missing data.
          2. Suppresses variable labels.
          3. Outputs Z-score file.
          4. Prints a reference dictionary.

```

```

STATISTICS: 1. Mean.
             2. Standard error.
             3-4. Not used.
             5. Standard deviation.
             6. Variance.
             7. Kurtosis.
             8. Skewness.
             9. Range.
            10. Minimum.
            11. Maximum.
            12. Sum.
            13. Coefficient of variation.
            14. 95% C.I. for sample mean.

```

```

CROSSTABS      VARIABLES=varlist(low,high)/
                TABLES=varlist BY varlist BY...
                OR
                TABLES=varlist BY varlist BY...

```

Two-way and n-way joint frequency tables program.

```

CROSSTABS      VARIABLES=RACE(0,1) ITEM1 TO ITEM10/
                TABLES=RACE BY ITEM1 TO ITEM10

```

```

STATISTICS:    ALL

```

```

OPTIONS: 1. Includes missing data.
          2. Suppresses all labels.
          3. Deletes row percentages.
          4. Deletes column percentages.
          5. Deletes total percentages.
          6. Deletes value labels but prints
             variable labels.
          7. Includes missing values in tables,
             not in statistics.
          8. Prints row variables in reverse order.
          9. Prints an Index of the tables produced.

```

```

STATISTICS: 1. Chi-square / Fisher's exact test if cases
               <21/Yates' corrected Chi-square for all
               2 x 2 tables.
            2. Phi for 2x2 tables, Cramer's V for larger
               tables.
            3. Contingency coefficient.
            4. Lambda, symmetric and asymmetric.
            5. Uncertainty coefficient, symmetric and
               asymmetric.
            6. Kendall's tau b.
            7. Kendall's tau c.
            8. Gamma
            9. Somer's D, symmetric and asymmetric.
            10. Eta.
            11. Pearson's R.

```

```

DISCRIMINANT GROUPS= { varname } (high,low)/
                      { SUBFILES }

VARIABLES= varlist/
ANALYSIS= varlist [(level)].../
          { DIRECT
          { WILKS
          { MAHAL
          { MAXMINF
          { MINRESID
          { RAO
[ METHOD /TOLERANCE=value/
          MAXSTEPS=value/FIN=value/FOUT=value/
          PIN=value/POUT=value/FOUT=value/VIN=value/
          FUNCTIONS=ng,ip,sig/
          PRIORS= { EQUAL
                  { SIZE
                  { value list } ]

```

Discriminant analysis program

```

RUN SUBFILES (A,B,C) (D,E)
DISCRIMINANT GROUPS=SUBFILES/VARIABLES=V1 TO V20/
              ANALYSIS=V1 TO V20/METHOD=RAO

```

```

OPTIONS: 10,13,16
STATISTICS: 5

```

- OPTIONS:
1. Includes missing data.
 2. Includes cases with missing values during classification.
 3. Suppresses step-by step output.
 4. Suppresses stepwise summary table.
 5. Prints classification results table.
 6. Prints discriminant scores and classification information.
 7. Prints a single plot of cases.
 8. Prints a separate plot for each group.
 9. Suppresses classification phase features for cases initially unclassified.
 10. Prints territorial map.
 11. Prints unstandardized discriminant function coefficients.
 12. Prints classification functions.
 13. Rotates the discrimination functions.
 14. Uses individual group covariance matrices for classification.
 15. Outputs matrix materials file.
 16. Allows matrix input.
 17. Outputs discriminant scores file.
 18. Outputs file of membership probabilities for all groups.
 19. Outputs file of actual group and classified group numbers.

- STATISTICS:
1. Means.
 2. Standard deviations.
 3. Pooled within groups covariance matrix.
 4. Pooled within groups correlation matrix.
 5. Matrix of pairwise F ratios.
 6. Univariate F ratios.
 7. Test for equality of group covariance matrices.
 8. Group covariance matrices.
 9. Total covariance matrix.

```

FACTOR          VARIABLES=varlist /
                { PA1
                { PA2
                { RAO
[TYPE = { ALPHA
        { IMAGE
        { BYPASS } /
DIAGONAL= v1,v2,../

```

SPSS - FREQUENCIES

```
NFACTORS= value /
MINEIGEN= value /
ITERATE=value /
STOPFACT=value /
ROTATE= {
  VARIMAX
  QUARTIMAX
  EQUIMAX
  OBLIQUE
  NOROTATE
}
DELTA= value /
FACSCORE[=mdrp ]/...]
```

Factor analysis program.

```
FACTOR      VARIABLES=VAR01 TO VAR30/
            TYPE=PA1/ROTATE=EQUIMAX/...
```

```
OPTIONS:    2
STATISTICS: ALL
```

- ```
OPTIONS: 1. Includes missing data.
 2. Deletes missing data, pairwise.
 3. Allows correlation matrix input.
 4. Allows factor matrix and communalities input.
 5. Outputs simple correlation matrix.
 6. Output factor matrix and communalities file.
 7. Outputs factor-score coefficient matrix file.
 8. Outputs means and standard deviations.
 9. Uses VARIABLE LIST with matrix input.
 (with option 3)
 10. Produces weighted factor scores when missing
 data is encountered. (with option 2 and 7)
 11. Sequence the output factor scores (with option7)
 12. Reads in mean, s.d. with correlation matrix.
 13. Output factor scores.
 14-23. NOT USED.
 24. Outputs fator scores in (4E20.14) format
 (unsequenced, i.e. option 11 not specified)
 or in (20X,3E20.14) format(sequenced,i.e.
 option 11 specified.
```

- ```
STATISTICS: 1. Means and standard deviations.
             2. Correlation matrix.
             3. Inverse and determinant of correlation matrix.
             4. Communalities, eigenvalues and proportion of
                total and common variance.
             5. Initial-factor matrix.
             6. Rotated-factor matrix and transformation matrix.
             7. Factor-scores coefficient matrix.
             8. Plot of rotated factors (with orthogonal
                rotations.)
```

```
-----
FREQUENCIES  GENERAL= {varlist}
                   {ALL}
```

```
INTEGER=varlist(low,high)[/varlist.../...]
```

One-way frequency distributions with descriptive statistics.

```
FREQUENCIES  INTEGER=NHHELP(0,7)INCOME TO EDUC INCOMER
              (0,99)
              GENERAL=INCOME, AGE, RACE, SEX, ITEM1 TO ITEM10
```

- ```
OPTIONS: 1. Includes missing data.
 2. Deletes value labels.
 3. Prints tables in 8-1/2 x 11 format.
 4. Outputs a permanent print file.
 5. Prints all tables in condensed format.
 6. Prints all tables greater than a page in
 condensed format.
 7. Deletes frequency tables.
```

8. Prints histograms.
  9. Prints a reference dictionary.
- STATISTICS:
1. Mean.
  2. Standard error.
  3. Median.
  4. Mode.
  5. Standard deviation.
  6. Variance.
  7. Kurtosis.
  8. Skewness.
  9. Range.
  10. Minimum.
  11. Maximum.
  12. Sum.
  13. Coefficient of Variation.
  14. 95% C.I. for sample mean.

---

G3SLS [CONSTANT variable name / ]  
[MATRIX matrix-identifier [=]value list /]  
[STORE store-identifier [=] value list /]  
[SYMMETRIC symmetric-identifier [=] value list /]  
[RESTRICT restrict-identifier [=] value  
{ store-identifier } { store-identifier } /  
or { or }  
{ matrix-identifier } { matrix-identifier }  
or { }  
{ constraints-identifier } { } ]

[CONSTRAINTS constraints-identifier [=] constraints-list/  
[HYPOTHESIS hypothesis-identifier [=] hypothesis-list/  
[MIXED mixed-identifier [=] value  
{ store-identifier } { store-identifier } { store-identifier }  
or { or }  
{ matrix-identifier } { matrix-identifier } { matrix-identifier }  
or { or }  
{ constraints-identifier } { symmetric-identifier }  
EQUATION equation-identifier [=] variable name  
[WITH] variable list /

MODEL [name ] [=] EQUATION [=] equation-list  
IDENTITY [=] equation list  
or  
ENDOG [=] variable list EXOG[=] variable list  
MIXED [=] mixed-list [RESTRICT [=] restrict-list]  
{ FTEST [=] restrict-list } [NO3SLS] [RESIDCOV] [COEFCOV]  
or  
RMONLY }

[CONSTRAINTS [=] constraints-list]  
[HYPOTHESIS [=] hypothesis-list] /

Generalized and three-stage least squares.

---

GUTTMAN SCALE scale name=varname(division point)...

Guttman scale analysis.

GUTTMAN SCALE MEMBSCAL=ALTMEM(1)  
OCCMEM(1) RECMEM(1)

STATISTICS: ALL

- OPTIONS:
1. Includes missing data.
  2. Suppresses variable labels.
  3. Suppresses ordering of variables.

- STATISTICS:
1. Inter-item and part-whole correlation coefficient.
  2. Coefficient of reproducibility.
  3. Minimum marginal reproducibility.
  4. Percent improvement achieved by Guttman scale.

## 5. Coefficient of scalability.

```

JFACTOR VARIABLES= varlist /
 [TYPE= { ULS
 GLS
 ML } /
 DIAGONAL=v1,v2,.. /
 NFACTORS=value /
 MINEIGEN=value /
 ITERATE=value /
 STOPFACT=value /
 ROTATE= { VARIMAX
 QUARTIMAX
 EQUIMAX
 OBLIQUE
 NOROTATE } /
 DELTA=value
 FACSCORE=[mdrp]/

```

Joresbog factor analysis program.

```

JFACTOR VARIABLES=VAR01 TO VAR30/
 TYPE=ULS/ROTATE=EUQEMAX/

```

```

OPTIONS: 2
STATISTICS: ALL

```

```

OPTIONS: 1. Includes missing data.
 2. Deletes missing data,pairwise.
 3. Allows correlation matrix input.
 4. Allows factor matrix and communalities
 input.
 5. Outputs simple correlation matrix.
 6. Output factor matrix and communalities file.
 7. Outputs factor-score coefficient matrix file.
 8. Outputs means and standard deviations.
 10. Print normalized residual correlation matrix.

```

```

STATISTICS: 1. Means and S.D.
 2. Correlation matrix.
 3. Inverse and determinant of correlation
 matrix.
 4. Communalities, eigenvalues and proportion
 of total and common variance.
 5. Initial-factor matrix.
 6. Rotated-factor matrix and transformation
 matrix.
 7. Factor-score coefficient matrix.
 8. Plot of rotated factors(with orthogonal
 rotations.)
 10. Print normalized residual correlation matrix.

```

MANOVA <dep var list> BY<factor list> WITH <covariate list> /

TRANSFORM(variable list) = { POLYNOMIAL (metric) /  
REPEATED/  
SPECIAL (matrix) / }

RENAME = (newname1,newname2 ...)/

ANALYSIS =<dep var list > WITH <covar list >/

PARTITION (factorname) = (df1,df2, ...)/

CONTRAST(factorname) = { DEVIATION/  
DIFFERENCE/  
SIMPLE/  
POLYNOMIAL(metric)/  
SPECIAL(matrix)/ }

SIGLEVEL = (significance level) /

PROTATE  
PROTATE(n)  
PROTATE(MAXEIGEN=m) = { NOROTATE/  
VARIMAX/  
EQUIMAX/  
QUARTIMAX/ }

CROTATE = { NOROTATE/  
VARIMAX/  
EQUIMAX/  
QUARTIMAX/ }

DESIGN = <list of effects in design>/

Multivariate Analysis of Variances

MANOVA WEIGHT1 BY SEX(1,2),DRUG(1,3)/  
PARTITION(DRUG)=(1,1)/  
CONTRAST(DRUG)=POLYNOMIAL(1,2,3)/  
DESIGN=SEX,DRUG(1),DRUG(2),SEX BY DRUG(1)  
+ SEX BY DRUG(2)

- OPTIONS:
1. Include missing data.
  2. Print parameter estimates only.
  3. Obtain last effect in design as residual.
  4. Omit grand mean from model.
  5. Print estimates of parameters for summary tables only.
  6. Print factor's last parameter as "negative sum".
  7. Input of with-cells variance-covariance matrix, cell means, and cell counts.
  8. Output of within-cells variance-covariance matrix, cell means, and cell counts.
  9. Input of within-cells correlation matrix, standard deviations, cell means, and cell counts.
  10. Output of within-cells correlation matrix, standard deviation, cell means, and cell counts.
  11. Output of error correlation matrices.
  12. Output of error variance-covariance matrices.
  13. Suppress special orthogonal processing.
  14. Perform computations using observations model.
  15. Calculate inner products in double precision.
  16. Suppress printing of parameter estimates.
  17. Print minimal multivariate output.

- STATISTICS:
1. Print observed cell means, deviations, and counts.
  2. Print variances-covariances and correlations of the estimated parameters.
  3. Print design matrix.
  4. Print error correlations.
  5. Print hypothesis sum-of-squares and cross-products.

SPSS - MULTIPLE RESPONSE

6. Print step-down F-tests table.
7. Print averaged F-test table.
8. Print raw discriminant function coefficients.
9. Print standardized discriminant function coefficients.
10. Print estimates of effects for canonical variates.
11. Print canonical correlations.
12. Print tests for homogeneity of variances.
13. Print test for homogeneity of variance-covariance matrices.
14. Print sum-of-squares and cross-products matrix for each cell.
15. Print variance-covariance matrix for each cell.
16. Print correlation matrix for each cell.
17. Print principal components of error correlation.
18. Print principal components of error covariance matrix.

MULT RESPONSE

GROUPS=group name [group label] (variable list (value list)) group name [group label] (variable list (value list)).../  
 VARIABLES=variable list (value list) variable list.../  
 FREQUENCIES=item list/  
 TABLES=item list BY item list [BY item list...]/  
 item list BY item list .../

MULT RESPONSE

GROUPS=PROBS NATIONAL PROBLEMS MENTIONED (PROB1 PROB2 PROB3(1,9))

MAGROR READERSHIP OF NATIONAL MAGAZINES (TIME REPUBLIC NEWSWEEK STONE (2))/  
 VARIABLES = EDUC(1,3)/  
 FREQUENCIES = PROBS MAGROR/  
 TABLES = PROBS BY EDUC MAGROR

OPTIONS:

1. Include missing values.
2. Causes the 'groupwise' deletion of missing values for multiple dichotomies.
3. Causes 'groupwise' deletion off missing values for multiple response groups.
4. Causes suppression of value labels for other than multiple dichotomies.
5. Causes the percents and totals in crosstab tables to be based on responses rather than on respondents.
6. Use 8Q-column on-line output format.
11. Forces a "three-up" condensed format for frequency tables involving other than a multiple dichotomy.
12. Causes the condensed format for frequency tables to be used for all items other than multiple dichotomies.

STATISTICS:

1. Include row percents in crosstab tables.
2. Include column percents in crosstab tables.
3. Include 2-way table total percents in cross-tab tables.

NONLINEAR

VARIABLES =dependent variable list WITH indep. list,  
 NB= no. of coefficients.  
 YHAT= arithmetic expression  
 G(I)= arithmetic expression ]  
 [B(1)= value,... B(NB)=value]  
 [BL(1)=value,... BL(NB)=value]  
 [BU(1)=value,... BU(NB)=value]

MODEL  
 [DERIVATIVES  
 PARAMETERS

[BNAME(1)=value,.. BNAME(NB)=value]  
[FIX(1)=value,... FIX(NB)=value]

{ MAXIT  
TOL1  
TOL2  
TOL3  
TOL4 } =parameter value

### Nonlinear regression

NONLINEAR VARIABLES= y WITH x, NB=3  
MODEL YHAT= B(1) + B(2) / (x+B(3))  
DERIVATIVES G(1)=1.  
G(2)=1./(x+B(3))  
G(3)= -B(2)/(x+B(3))\*\*2  
PARAMETERS B(1)=0.  
B(2)=10.  
B(3)=2.  
OPTIONS 3  
STATISTICS 4,7,8

OPTIONS: 1. Include missing data.  
2. Use Gauss' method in performing regression.  
3. Suppress derivative checking.  
4. Output formatted for 80-columns per line.

STATISTICS: 1. Print derivative checking report.  
2. Print singular values and vectors of Z Z  
at the initial values of the parameters and  
at the final parameter values.  
3. Print a detailed iteration description.  
4. Print iteration summary.  
5. Print approximate variance-covariance matrix  
of parameters.  
6. Print normalizing elements and correlation  
matrix of parameters.  
7. Print 95% C.I. about the parameters assuming  
a linear hypothesis.  
8. Print the results of exploration.  
9. Print estimated values of dependent variable  
and residual table.

---

NONPAR CORR {varlist WITH varlist}  
{varlist}

Spearman and/or Kendall rank-order correlation  
coefficients.

NONPAR CORR RESDYTH, INCOME, NACT, EDRESPON, OCLEURES

OPTIONS: 1. Includes missing data.  
2. Deletes missing data, listwise.  
3. Two-tailed test of significance.  
4. Outputs correlation matrix file.  
5. Kendall correlations.  
6. Kendall and Spearman correlations.

STATISTICS: none, no additional statistics available.

---

NPAR TESTS test name [parameter]=varlist [parameters] / test name.  
ALL

BINOMIAL = varlist {cutting point  
or  
value 1, value 2}

CHI-SQUARE= varlist[(low,high)] /  
[EXPECTED= fi...fn/]

K-S {UNIFORM  
NORMAL  
POISSON} [parameter]=varlist /

```

RUNS { MEAN
 { MEDIAN
 { MODE
 { value } =varlist

MCNEMAR =varlist [WITH varlist]/
SIGN =varlist [WITH varlist]/
WILCOXON =varlist [WITH varlist]/
COCHRAN =varlist/
FRIEDMAN =varlist/
KENDALL =varlist/
MEDIAN =(value) = dep. varlist BY
 indep. variable
 (value 1,value 2)

```

```

{ M-W
 { K-S
 { W-W
 { MOSES
 { K-W } =dep. varlist BY indep. variable
 (value 1, value 2)/

```

Non-parametric statistics.

NPART TESTS K-S= INCOME BY SOCSCL(1,2)/

- OPTIONS:
1. Include missing data.
  2. Listwise deletion of missing data.
  3. Special handling of variable lists for paired variable tests.
  4. Use random sampling if not all cases will fit into the maximum memory available.
  5. Output will be restricted to 75 printed columns.
  6. Suppress variable labels.

STATISTICS: 1. Mean, maximum, minimum, SD, no. of valid cases for mentioned variables.

ONEWAY dependent varlist BY independent variable (min,max)/

POLYNOMIALS=n/CONTRAST=n1,n2.../  
 RANGES= testname (alpha) where testname=

```

{ LSD
 { DUNCAN
 { SNK
 { TUKEYB
 { TUKEY
 { LSDMOD
 { SCHEFFE

```

One-way analysis of variance.

ONEWAY DEPVAR BY INDEPVAR(1,6)/  
 CONTRAST=1,-.5,-.5/  
 POLYNOMIAL=3/  
 RANGE5=LSD(.10)

STATISTICS: ALL

- OPTIONS:
1. Includes missing data.
  2. Deletes missing data, listwise.
  3. Suppresses variable labels.
  4. Outputs a file with number of cases, mean, and standard deviation for each category.
  5. Not used.
  6. Uses first 8 characters of value labels of independent variable as group labels.
  7. Uses input of category frequencies, means and standard deviations, not raw data.
  8. Uses input of category frequencies, means, pooled variance and degrees of freedom for the pooled variance.
  10. Uses harmonic mean for ALL group in range test.

STATISTICS: 1. Number of cases for each category, means, standard deviations, standard errors, minimum, maximum and 95% confidence interval for the mean.  
 2. Both fixed and random effects measures.



3. Homogeneity of variance statistics.
4. Print random effects (MODEL II) estimate of the between component variance.

PARTIAL CORR varlist WITH varlist BY varlist(order values)/...  
matrix type varlist BY varlist(order values)/...

Partial correlation coefficient program.

PARTIEL CORR EDUCATN WITH INCOME BY FATHRED, FATHRINC(1)

- OPTIONS:
1. Includes missing data.
  2. Deletes missing data, pairwise.
  3. Two-tailed test of significance.
  4. Allows matrix input.
  5. Outputs simple correlation matrix.
  6. Uses VARIABLE LIST with matrix input.  
(with option 4)
  7. Omits degrees of freedom with significance.
  8. Prints only nonredundant partials.
  20. Substitutes mean for missing data when computing the correlation matrix.

- STATISTICS:
1. Zero-order correlations.
  2. Means and standard deviations.
  3. Correlation matrix only if non-computable correlations are encountered.

PEARSON CORR varlist WITH varlist  
varlist

Pearson product moment correlations for pairs of variables.

PEARSON CORR POP60 WITH MEDSCH TO PTTERTRY

OPTIONS: 3  
STATISTICS: ALL

- OPTIONS:
1. Includes missing data.
  2. Deletes missing data, listwise.
  3. Two-tailed test of significance.
  4. Outputs correlation matrix file.
  5. Omits number of cases and significance.
  6. Prints only nonredundant coefficients.
  7. Compute correlation matrix using temporary means.
  20. Substitute mean for missing data when computing correlation matrix.

- STATISTICE:
1. Means and standard deviations.
  2. Cross product deviations and covariance.

REGRESSION  
[METHOD= { BACKWARD  
FORWARD  
STEPWISE } ]  
VARIABLES= varlist/  
REGRESSION= dependent variable  
(NSTEPS, FIN, TOL, FOUT) WITH  
independent variables(inclusion level)/  
[RESIDUALS=mdrp ]

Stepwise multiple linear regression program.

OPTIONS: 2,4  
STATISTICS: ALL

- OPTIONS:
1. Includes missing data.
  2. Deletes missing data, pairwise.
  3. Suppresses variable labels.
  4. Allows matrix input.

# SPSS - RELIABILITY

5. Means and standard deviations with matrix input. (with option 4)
6. Suppresses step-by-step output.
7. Suppresses the summary table.
8. Outputs correlation matrix file.
9. Uses VARIABLE LIST with matrix input. (with option 4).
10. Sequence residuals file.
11. Output of Residuals.
12. Write y' and residuals on alternative output file.
13. Suppress the plotting of cases with missing data.
14. Suppresses printing of axes on plots of residual vs. y'
15. Output of means and standard deviations.
16. Variance/covariance matrix output.
17. Output normalized regression coefficients (Bs).
18. Output unnormalized regression coefficients (Bs).
19. Force regression through the origin.
20. Substitute mean for missing data.
21. Print t, not F in step-by-step output.

## STATISTICS:

1. Correlation matrices.
2. Means, standard deviations and number of cases.
3. Correlation matrix only if non-computable correlations are encountered.
4. Plot residuals against sequence of cases in a file.
5. Statistics for residuals Von Neumann Ratio, Durbin-Watson statistics and a runs test are computed.
6. Plot standardized residuals against standardized y' values.
7. Print correlation matrix with number of cases for each coefficient.
8. Print estimated variance-covariance matrix of the unnormalized regression coefficients.
9. Compute 95% C.I. for unnormalized regression coefficient (B).
10. Compute 95% C.I. for normalized regression coefficients. (BETA).

---

## RELIABILITY

```
VARIABLES= varlist/
SCALE(label) = scale list /
MODELS= { ALPHA
 SPLIT
 GUTTMAN
 STRICTPARALLEL
 PARALLEL }
 or
FORMATS = INPUT(i/j), OUTPUT(k/l)
VARIABLES=Varlist /...
```

### Item and scale analysis program

- OPTIONS:
1. Includes missing data.
  2. Not used.
  3. Suppresses labels.
  4. Allows input of covariance matrix.
  5. Allows input of standard deviations and correlation matrix.
  6. Allows input matrix to be triangular.
  7. Allows input of means.
  8. Outputs covariance matrix file.
  9. Uses VARIABLE LIST with correlation matrix input.
  10. Stops after writing a covariance matrix on BCD out.
  11. Outputs item means on BCD out.
  12. Outputs triangular matrices as single

- vectors.
- 13. Outputs matrices as single vectors rather than by rows.
- 14. Space saver solution.
- 15. Automatic deletion of items with zero variance.
- 16. Continue processing with zero variance items.
- 17. Friedman's X .
- 18. Cochran's Q.

- STATISTICS:
- 1. Item means, standard deviations.
  - 2. Covariance matrix.
  - 3. Correlation matrix.
  - 4. Scale mean, scale variance.
  - 5. Mean, SD, minimum, maximum and range of item means.
  - 6. Mean, SD, minimum, maximum and range of item variance.
  - 7. Mean, SD, minimum, maximum and range of interitem covariance.
  - 8. Mean, SD, minimum, maximum and range of interitem correlation.
  - 9. Prints total statistics.
  - 10. Analysis of variance.
  - 11. Tuckey test for additivity.
  - 12. Hotelling T-squared.

SCATTERGRAM

```
varlist(low,high)
varlist(LOWEST,high)
varlist (low,HIGHEST)
varlist(low,high) WITH varlist(1,h)
```

Two-way scattergram plots of pairs of variables.

SCATTERGRAM VAR003(0,10) WITH VAR006 (7,HIGHEST)  
VAR007 TO VAR009(LOWEST,99)

- OPTIONS:
- 1. Includes missing data.
  - 2. Deletes missing data, listwise.
  - 3. Suppresses variable labels.
  - 4. Suppresses plot grid lines.
  - 5. Prints diagonal grids.
  - 6. A two-tailed test of significance. (with statistic 3)
  - 7. Automatic scaling.
  - 8. If not enough core is available, plots of the first n cases are produced.

- STATISTICS:
- 1. Pearson r.
  - 2. r-squared.
  - 3. Significance of r.
  - 4. Standard error of the estimate.
  - 5. Intercept with the vertical axis.
  - 6. Slope.

SPECTRAL

```
{ UNIVARIATE }
{ BIVARIATE } =varlist /

WINDOW = { PARZEN
 BARTLETT
 TUKEY
 HAMMING
 UNIT } [(lags)] /

LAGS = <lag> /
 <unit> [(period)]
```

Spectral Analysis of Time Series

SPSS - SUMMARY TABLES

SPECTRAL      BIVARIATE= VAR1,VAR2 /  
                   WINDOW = HAMMING(10) /

OPTIONS        4,5,18

STATISTICS    ALL

- OPTIONS:
1. Ignore missing value indicators.
  2. Listwise replacement of missing observations by the mean.
  3. Suppress variable labels.
  4. Plot auto-correlation and spectrum for each variable.
  5. Plot unweighted cross-correlation, phase and coherence.
  6. Plot weighted cross-correlation, phase and coherence.
  7. Not used.
  8. Not used.
  9. Write unweighted auto-covariances to alternate output file.
  10. Write weighted auto-covariances to alternate output file.
  11. Write spectrum auto-covariances to alternate output file.
  12. Write unweighted cross-covariances to alternate output file.
  13. Write weighted cross-covariances to alternate output file.
  14. Write co-spectrum and quadrature spectrum to alternate output file.
  15. Write phase and coherence to alternate output file.
  16. Write phase and amplitude to alternate output file.
  17. Write gains to alternate output file.
  18. Use the unbiased estimator of the covariance, rather than the positive-definite estimator.

Only one "write" option may be selected from the univariate set, 9, 10, 11; similarly, only one from the bivariate set, 12 through 17.  
 Each statistic requested is produced for each window-lag pair.

- STATISTICS:
1. For each variable, at each lag/band  $j$ :  
 $j=0, 1, \dots, m$   
 auto-covariance,  
 weighted covariance,  
 auto-correlation,  
 weighted correlation,  
 power spectrum,  
 spectral density,  
 log<sub>10</sub> of the power spectrum,  
 period of the frequency band.
  2. For each variable:  
 standardized bandwidth of the window,  
 equivalent degrees of freedom,  
 variance ratio  
 variance.
  3. For each bivariate pair(x,y), at each lag/band  $j$ :  
 cross covariance of x with lagged y,  
 weighted covariance of x with lagged y,  
 cross covariance of y with lagged x,  
 weighted covariance of y with lagged x,  
 co-spectrum,  
 quadrature spectrum.
  4. Not used.
  5. For each bivariate pair, (x,y), at each band  $j$ :  
 gain from x to y,  
 gain from y to x,  
 phase shift,  
 amplitude,  
 coherence,

period of the band,  
frequency of the band.

SUMMARY TABLES

dependent variable list [(label)] BY  
independent varlist (min,max)  
[BY control variable list (min,max BY...)]

Combines several crosstabulation of dichotomous variables  
into one single table.

SUMMARY TABLES      RULE1 TO RULE9 BY SEX(1,2)

- OPTIONS:
1. Ignore missing data indicator.
  2. Listwise deletion of missing data.
  3. Suppress printing of variable labels for dependent variables.
  4. Suppress printing of value labels for independent and control variables.
  5. Suppress printing of both variable labels and value labels for independent and control variables.
  6. Print the count of non-zero cases in each cell.
  7. Cause percentages to be computed within rows instead of within categories of independent variable.

STATISTICS:      no statistics are available.

SURVIVAL

survival time variable BY group variable (lo,hi)/  
INTERVALS = THRU upper limit BY interval width /  
[PLOT = (group1, group2,...groupn ).../]  
[COMPARE = (group1, group2, groupn) .../]

Survival Analysis

SURVIVAL              ADMSURV BY RACE(1,3)/  
                         INTERVALS= THRU 98 BY 2/  
                         PLOT =(1,2,3)/ COMPARE=(1,2,3)

- OPTIONS:
1. Ignores missing data indicator.
  2. Causes pairwise comparison of each of the groups named in a COMPARE specification.
  3. Causes only comparisons between groups to be performed.

STATISTICS:      there are currently no optional statistics  
                         available with SURVIVAL.

TETRACHORIC

VARIABLES=varlist (low,high)/  
CORRELATIONS=varlist WITH varlist/

Tetrachoric correlation coefficients

TETRACHORIC VARIABLES=VAR001 TO VAR009(1,2)/  
                         CORRELATION=VAR001 TO  
                         VAR008 WITH VAR009/

- OPTIONS:
1. Includes missing data.
  2. Deletes missing data listwise.
  3. Suppresses variable labels.
  4. Invokes special handling for paired variable lists.
  5. Causes Hotelling's T\*\*2 statistic for the equality of two vectors of means to be computed and printed.

STATISTICS:      none. No additional statistics available.

SPSS EXAMPLES=====

## SPSS - EXAMPLES

These examples are oriented to the CDC NOS operating system.

-----  
 An example of an SPSS job in which the data is on punched cards and only one procedure (viz. "FREQUENCIES") operates on that data:

```
GRAB(SPSS)
SPSS.
7/8/9 (END-OF-RECORD CARD)
RUN NAME COMPLETE DECK FOR RUN FROM CARDS
VARIABLE LIST AGE,SEX,RACE,INCOME,EDUCATN
INPUT FORMAT FIXED(5X,F2.0,A1,3F1.0)
RECODE SEX('M'=1) ('F'=2) (ELSE=0)
VALUE LABELS SEX (1) MALE (2) FEMALE/
 RACE (1) WHITES (2) BLACK (3) INDIANS
MISSING VALUES
INPUT MEDIUM CARD
N OF CASES 10
FREQUENCIES GENERAL=ALL
OPTIONS 1,4
STATISTICS ALL
READ INPUT DATA
 20F685
 40M266
 .
 (10 DATA CARDS IN ALL)
 .
FINISH
6/7/8/9 (END-OF-FILE CARD)

```

An example of an SPSS job in which the data is on punched cards where the number of cases is unknown and more than one procedure (eg. "CONDESCRIPTIVE", & "CROSSTABS") is used on that data:

```
GRAB(SPSS)
SPSS.
7/8/9 (END-OF-RECORD CARD)
RUN NAME COMPLETE DECK FOR RUN FROM CARDS
FILE NAME EXAMPLE
DATA LIST FIXED(2)/1 AGE6-7/2 SEX 8(A),
 RACE,INCOME,EDUCATN 9-11
INPUT MEDIUM CARD
N OF CASES UNKNOWN
CONDESCRIPTIVE ALL
OPTIONS 3,4
STATISTICS ALL
READ INPUT DATA
 .
 (UNKNOWN NUMBER OF CASES)
 .
7/8/9 (END-OF-RECORD)
COMMENT END-OF-RECORD CARD TO SIGNAL END OF DATA
*RECODE SEX('M'=1) ('F'=2) (ELSE=0)
CROSSTABS TABLES=AGE BY SEX TO EDUCATN
STATISTICS ALL
FINISH
6/7/8/9 (END-OF-FILE)

```

An example of an SPSS job in which the input data is read from the permanent disc file "SPSSDATA":

```
GET(DISK,SPSSDATA)
GRAB(SPSS)
SPSS.
7/8/9 (END-OF-RECORD)
RUN NAME DATA ON DISC RATHER THAN ON CARDS
FILE NAME EXAMPLE
VARIABLE LIST VAR01 TO VAR05
INPUT FORMAT FREEFIELD

```

```

COMMENT 'FREEFIELD' MEANS THE NUMBER MAY
COMMENT BE IN ANY FORMAT AS LONG AS THEY ARE
COMMENT SEPARATED BY BLANKS AND ORDERED WITH
COMMENT RESPECT TO VARIABLE LIST.
VAR LABELS VAR01,AGE/VAR02,SEX/VAR03,RACE/VAR04
 INCOME/VAR05 EDUCATION
VALUE LABELS VAR02 (1) MALE (2) FEMALE
INPUT MEDIUM DISK
N OF CASES 100
LIST CASES CASES=10/VARIABLES=VAR01 TO VAR05
COMMENT TO LIST FIRST 10 CASES OF THE
COMMENT SUBFILE GROUP
READ INPUT DATA
SCATTERGRAM VAR01 TO VAR05
FINISH
6/7/8/9 (END-OF-FILE)

```

---

An example of an SPSS job in which a "save" file (i.e. SPSS control cards and data) is created and stored in a permanent disc file "MYDATA":

```

GRAB(SPSS)
SPSS.
SAVE(SVFILE,MYDATA)
7/8/9 (END-OF-RECORD)
RUN NAME SAVE FILE RUN
FILE NAME EXAMPLE
VARIABLE LIST AGE,SEX,RACE,INCOME,EDUCATN
.
. (OTHER SPSS COMMANDS)
.
READ INPUT DATA
.
. (DATA DECK)
.
COMMENT ANOTHER SPSS PROCEDURE AFTER THE DATA
T-TEST GROUPS=SEX(1)/VARIABLES=INCOME
SAVE FILE
FINISH
6/7/8/9 (END-OF-FILE)

```

---

An example of an SPSS job in which a "SAVE" file is retrieved from the permanent disc file and used as the data from a "CROSSTABS" procedure.

```

GET(GTFILE,MYDATA)
COMMENT THE LOGICAL FILE NAME ON THE GET CARD MUST BE
COMMENT "GTFILE"
GRAB(SPSS)
SPSS.
7/8/9 (END-OF-RECORD)
RUN NAME USING A SAVE FILE
GET FILE EXAMPLE
COMMENT THE PARAMETER IN THE SPECIFICATION FIELD OF THE
COMMENT "GET FILE" CARD MUST BE THE SAME AS THE PARAMETER
COMMENT USED IN THE SPECIFICATION FIELD OF THE FILE NAME
COMMENT CARD WHEN THE FILE WAS CREATED.
CROSSTABS AGE BY SEX BY RACE
.
. (OTHER SPSS COMMANDS)
.
FINISH
6/7/8/9 (END-OF-FILE)

```

# BMDP - BIOMEDICAL COMPUTER PROGRAMS

## BMDP INTRODUCTION=====

The BMDP computer programs are general purpose and designed to handle a wide variety of statistical problems. To adapt the programs to a specific problem, parameters must be specified that describe the characteristics of the data to be analyzed, the mode of analysis desired and choice of output.

Important problem parameters include:

- problem title (identifies the problem)
- description of the input data (format, size and source)
- description of the variables (names and limits)
- transformations of variables
- modes of analysis (specific to the various problems)
- choice of output (numerical results and plots)

Variables can be given an alphanumeric name up to 8 characters in length. A name can contain embedded blanks if the name is specified in quotes.

Continuous variables can be categorized into intervals by specifying CUTPOINTS. Categories can be given names in the GROUP paragraph.

Data within BMDP programs on which statistical analysis is done are always processed as decimal values.

Format statements for BMDP use the FORTRAN F, X, A, / and repetition specifiers only.

## BMDP CONTROL LANGUAGE=====

THE BMDP control language is placed on input cards or records in free format, the only restrictions being that the control information for each new problem or subproblem must start on a new 80 column card or record.

### PARAMETER NAME

A parameter name identifies a control language paragraph or sentence. Specific parameter names have defined meanings as given in the program descriptions.

/PROBLEM TITLE='HEART STUDY'.

PROBLEM and TITLE are parameter names.

### SUBSCRIPTING

If parameters need to be used more than once in a paragraph in reference to different variables, subscripts may be used to identify the variable being referred to.

CUTPOINTS(5) = 1,3,7.  
CUTPOINTS(2) = 5,9,15.

These two sentences refer to variables 5 and 2 respectively.

-----  
\*BMDP-77, Biomedical Computer Programs, P-Series, Edited by W.J. Dixon and M.B. Brown, University of California Press, Berkeley, 1977.  
-----



## SENTENCES

Each control language command or parameter is specified in a sentence which must end with a period. Most sentences assign values.

```
CASES = 213.
FORMAT = '(5F7.0)'.
```

Sentences can be directive.

```
REWIND.
NO REWIND.
NO CORRELATION.
```

## PARAGRAPHS

Sentences are grouped into paragraphs. The name is its first word (eg. INPUT, PROBLEM, SAVE, etc.). A '/' must appear either before the paragraph name or at the end of a paragraph. Although each type of paragraph allows only specific sentences, the order of these allowable sentences in the paragraph is arbitrary (except the PROBLEM paragraph which must appear first).

```
/PROBLEM TITLE = 'HEART STUDY'.
/INPUT VARIABLES ARE 5.
 CASES = 213.
 FORMAT IS '(5F7.0)'.
/VARIABLE NAMES ARE 'WT IN 72', HEIGHT, AGE, 'PULSE RT', SEX.
/END.
.
.
 . 213 data cards
/FINISH.
```

## RESERVED WORDS

IS, ARE, TO and BY are reserved words and can only be used as variable names if placed in quotes.

## SPECIAL FEATURES

### BLANKS

Blanks can be used freely to space program control information, except in the middle of names or numbers.

### EQUAL SYMBOL, IS or ARE

The equal symbol '=' and the words 'is' and 'are' can be used interchangeably.

```
CASES = 213.
FORMAT IS '(5F7.0)'.
```

### REPETITION

Repetition of the same numerical values in a list can be specified by the symbol '\*'.  
1,2,3,3,3,4,6,6,6,6 is the same as 1,2,3\*3,4,4\*6

### IMPLIED LIST

To facilitate specification of a list of numbers which are equally spaced, the procedure:

beginning # TO ending # BY step  
may be used.

For example the list:  
1,2,3,4,5,5.5,6,6.5,7  
May be given by

# BMDP - MAIN CHART

1 TO 5 BY 1, 5.5 TO 7 BY .5

## TABBING

A number in parenthesis preceding a value can be used as a subscript of the list element that follows

NAMES = AGE, SEX, (15) WEIGHT, PULSE.  
give names to variables 1,2,15 and 16.

INPUT and FORMAT  
are the same as  
INPT and FRMT

## PARAGRAPH NAME REPETITION

The name of a paragraph can be repeated before each sentence if desired.

## COMPOUND ASSIGNMENT SENTENCES

Elements in lists of two or more assignment sentences in the same paragraph can be stated in a form to facilitate matching. For example:

VARIABLE NAMES = AGE, WEIGHT, SEX, CHANGE.  
MINIMUMS = 0,0,1,-10.  
MAXIMUMS = 20,275,2,10.

can be written:

| VARIABLE NAMES, | MINIMUMS, | MAXIMUMS = |
|-----------------|-----------|------------|
| AGE,            | 0         | , 20,      |
| WEIGHT,         | 0         | , 275,     |
| SEX,            | 1         | , 2,       |
| CHANGE,         | -10       | , 10.      |

repetition, implied list, subscripting or tabbing cannot be used with this feature.

## REPEAT DECLARATIONS

The same location in a list can be specified more than once, but only the last is retained.

CONTRAST = 4\*1, -7, 3\*1.

can be written

CONTRAST = 8\*1, (5), -7.

---

## MAIN CHART PARAGRAPH AND PARAMETER DECLARATIONS

The following table lists the paragraph and parameter names which are common to most of the BMDP programs. In the individual program descriptions, those names which have a function in that program identical to that described in this table will be mentioned in name only and the user can refer to this table for comments. If an option is specific to a program, it will be explained in that program. Default values are listed in parenthesis at the end of comments.

| PARAGRAPH NAME      | PARAMETER NAME | COMMENTS                                                                        |
|---------------------|----------------|---------------------------------------------------------------------------------|
| /PROB<br>(required) |                | Signifies the beginning of a new problem and is required to begin each problem. |
|                     | TITLE='text'.  | Text in quotes to describe the problem. Max. length 160 characters.(blank)      |

/INPUT  
(required on  
first problem)

VARIAB=#.

Input describes the input data size format and source. Required for the first problem of a deck setup.

States the number of variables required unless input is from a save file. (none/prev.)

VARIABLES ARE 10.

FORMAT = 'C'

Specifies read input format. Must be less than 800 characters long. Required unless input is a save file. (none/prev.)

FORMAT is '(5F7.0,3F10.2)'

CASE = #.

Specifies the number of cases to be read. By default, read stops when EOF is reached. (EOF/prev.)

CASES = 300.

UNIT = #.

Specifies logical unit number for data or save file. Must be stated unless the data is with the control cards. Unit number is 5 by default. Cannot be 1,2, or 6. Should be less than 10. (5/prev.)

UNIT is 9.

REWIND.

Specifies REWIND to rewind the input file. NO REWIND suppresses rewind. Input unit will rewind by default after input data is read. (YES/prev.)

NO REWIND.

CODE = C.

Required if input is from a save file. The code name (1-8 characters) must correspond to the code given to the save file when originally stored. (input controlled by format)

CODE IS MYFILE.

CONTENT = C.

Specifies what type of save file is desired (eg. DATA, CORR, MEAN OR FREQ). If not stated, the first subfile whose code matches the one specified will be read. (none)

CONTENT IS CORR.

LABEL = 'C'.

Allows identifying desired input if the same CONTENT (ie. DATA) is recorded more than once under the same code. A label can be up to 40 characters long. (none)

LABEL IS 'COMPUTED IN BMDP4M'.

TYPE = name.

Specifies the way the program should treat the input. Acceptable types vary from program to program. (DATA)

TYPE = CORR.

TYPE = FREQ.

# BMDP - MAIN CHART

/VARIAB  
(optional if input from  
BMDP save file)

Specifies variable characteristics.  
This paragraph is not required if  
all defaults are acceptable.

ADD = #.

Specifies the number of variables to  
be added by transformations. Can be  
negative if the transformation  
reduces the number of  
variables(counting from the right).  
(0 /prev.)

NAME = list.

Assigns 1-8 character names to the  
variables. Names must be enclosed  
in single quotes if they contain  
embedded blanks.Names are optional.

NAMES ARE HEIGHT, 'WT LBS'.

USE = list.

Specifies which variables are actu-  
ally going to be used in this pro-  
blem. Optional. (all variables)

USE = 1,3, 12 TO 16.

LABEL = list.

Specifies variables (up to two var-  
iables of 4 characters each) which  
are to act as case labels. (none  
/prev.)

LABELS are 4,5.

MISS = #list.

Missing value list for the vari-  
ables. (none/prev.)

MISSING ARE 5\*999.

GROUP = v.

Specifies a variable on which the  
data will be group.(none/prev.)

GROUPING is 3.

MAX = #list.

Maximum value list.(none/prev.)

MIN = #list.

Minimum value list.(none/prev.)

MAX = 35, 12, (15)8.  
MIN = 0, 0, (15)1.

BEFORET.or  
AFTERT.

Determines whether data checking  
(MISS, MAX & MIN) is to be done  
before or after transformations.  
(BEFORET/prev.)

/GROUP or  
/CATEG  
(optional)

Specifies group or category cut-  
points and associated name for  
variables. This paragraph is  
optional. If any cutpoints, codes  
or names are given, all from a  
previous subproblem are eliminated.  
If none are stated, any from a  
previous subproblem are still in  
effect.

CUTP (#) =  
#1,#2,....

Cutpoints for grouping or categor-  
izing a continuous variable. n-1  
cutpoints determine n categories.  
(CODE/prev.)

CODE (#) =  
#1,#2,....

For categorical variable, specifies  
the values which are to act as  
grouping values for the data set. If  
no codes are given for a grouping  
variable, cutpoints will be used.  
If neither given, the 10 smallest  
values are used.

NAME (#) =  
C1,C2,....

A 1-8 character name for cutpoints or codes. If not stated, cutpoint or code values will be used for names when necessary.

CUTPOINTS (5) ARE 2, 4.  
CODES (4) = 5,7,3 10.  
NAMES (5) = BABY, TODDLER,  
          OLDER.  
NAMES (4) = A, B, C, D.

RESET.

If RESET, all assignments in previous GROUP paragraph are reset to preassigned values. (not reset)

/TRANSF  
(optional)

Provides capability for minimal computation. A FORTRAN subroutine capability exists for more extensive transformations. If paragraph consists only of TRANSF, the transformations used in the preceding problem will still be in effect. If any new transformations are defined, all old ones are eliminated. Variable may be referenced for transformation by name or by index.

Up to 7 separated temporary variables may be used.

Allowed transformations

- Arithmetic

simple assignment  $X(1)=X(10)$ .  
binary operation  $X(1)=X(1)+X(2)$ .  
functions  $X(1)=SQRT(X(3))$ .

- Logical  $X(7)=X(1) \text{ LT } X(2)$ ., where  
 $X(7)$  is 1(true) if  $X(1)$  is less than  $X(2)$  and is 0(false) otherwise.

- Random number generation  
 $X(1)=RNDU(\text{starting value})$ .

Allowable forms

- result = argument  
- result = arg1 operator arg2  
- result = function (argument)

- where result can be:  
 $X(i)$  for  $i$  an integer  
a variable name  
a temporary variable  
a word USE

- and the argument can be:  
 $X(i)$  for  $i$  an integer  
a variable name  
a previously defined  
temporary variable  
the words USE, KASE or XMIS  
a numerical constant

- USE can be used to select cases for analysis.

USE =  $X(1) \text{ LT } X(2)$ .

- uses those cases for which  
 $X(1)$  is less than  $X(2)$ .

# BMDP - TRANSFORMATIONS

- KASE is the case sequence #.  
X(5) = KASE.
- places a sequential number in variable 5 which numbers the cases.
- XMIS is the internal missing value code.  
TEMP = X(1) EQ 7.  
X(1) = XMIS IF TEMP.
- states that the first variable is assigned the missing value whenever it has a value of 7.

/SAVE  
(optional)

Causes information (data, statistical results, etc.) to be written to a save file. (Not required)

UNIT = #.

Logical unit to which the save file is to be written. Cannot be 1, 2, 5 or 6. Should be less than 10. (none)

NEW.

Required if this is a new save file. (not new)

CODE = C.

A 1-8 character label for the save file. (none)

CONTENT =  
name.

Specifies what is to be written (ie. DATA, FREQ, CORR, MEAN, etc). Check individual programs for options available. (DATA)

LABEL =  
'text'.

Up to a 40 character label to identify specifically the save file. (blank)

```

/SAVE NEW.
 UNIT = 9.
 CONTENT IS DATA.
 CODE IS WERNER.
 LABEL = 'WERNER DATA AS OF
 3-13-74'.

```

/END  
(required)

Required as last paragraph.

/FINISH

Terminates BMDP. Last card in deck. BMDP programs dated before August, 1977, require the FINISH paragraph when the data are not on cards with the command deck or when the number of cases is specified. The FINISH paragraph must not be used when the number of cases is not specified and data is included on cards. The FINISH paragraph is not necessary in BMDP programs dated August 1977 or later.

## TRANSFORMATION SUMMARY TABLE

| KIND OF<br>TRANSFOR-<br>MATIONS | HOW<br>USED | RESULT<br>COMMENTS | TYPICAL<br>EXAMPLE |
|---------------------------------|-------------|--------------------|--------------------|
| -----                           | ----        | -----              | -----              |

|                                      |         |                                                                                                   |                       |
|--------------------------------------|---------|---------------------------------------------------------------------------------------------------|-----------------------|
| Arithmetic<br>(simple<br>assignment) | a=b     | b placed into a                                                                                   | X(7) = X(1).          |
| Arithmetic<br>(functions)            | LOG(a)  | log <sub>10</sub> (a)                                                                             | WEIGHT = LOG(WEIGHT). |
|                                      | SQRT(a) | square root of a                                                                                  | X(2) = SQRT(X(5)).    |
|                                      | EXP(a)  | to the power of a                                                                                 | X(1) = EXP(X(2)).     |
|                                      | ABS(a)  | a , absolute<br>value of a                                                                        | CHANGE = ABS(X(17)).  |
|                                      | SIN(a)  | trig sine of a                                                                                    | X(1) = SIN(X(1)).     |
|                                      | COS(a)  | trig cosine of a                                                                                  | X(2) = COS(X(2)).     |
|                                      | ATAN(a) | trig arc tangent<br>of a                                                                          | X(3) = ATAN(X(3)).    |
|                                      | INT(a)  | integer part of a                                                                                 | AGE = INT(AGE).       |
|                                      | SIGN(a) | $\begin{cases} -1 & \text{if } a < 0 \\ 0 & \text{if } a = 0 \\ 1 & \text{if } a > 0 \end{cases}$ | X(5) = SIGN(X(1)).    |
|                                      | NONB(a) | blank field<br>detection returns<br>missing value if<br>a is blank                                | X(7) = NONB(X(7)).    |
| Arithmetic<br>(binary)               | a+b     | addition                                                                                          | X(3) = X(1)+X(2).     |
|                                      | a-b     | subtraction                                                                                       | WEIGHTS=WEIGHT-X(3).  |
|                                      | a*b     | multiplication                                                                                    | DIST=RATE*TIME.       |
|                                      | a/b     | division                                                                                          | X(1)=1.0/TIME.        |
|                                      | a**b    | exponentiation                                                                                    | X(1)=HEIGHT** 3.      |
|                                      | aMODb   | the remainder<br>of a divided<br>by b                                                             | MONTH=AGE MOD 12.     |
| Logical<br>Operation                 | a LE b  | 1 if a<b, 0 if a>b                                                                                | USE = X(1) LE X(2).   |
|                                      | a LT b  | 1 if a<b, 0 if a>b                                                                                | S(1) = X(2) LT X(3).  |
|                                      | a GE b  | 1 if a>b, 0 if a<b                                                                                | USE = WEIGHT GE 50.   |
|                                      | a GT b  | 1 if a>b, 0 if a<b                                                                                | T = W GT 25.          |
|                                      | a NE b  | 1 if a≠b, 0 if a=b                                                                                | USE = X(2) NE X(3).   |
|                                      | a EQ b  | 1 if a=b, 0 if a≠b                                                                                | X(1) = X(2) EQ X(3).  |
|                                      | a AND b | 1 if a≠0 and b≠0                                                                                  | USE = TEMPA AND T.    |
|                                      | a OR b  | a if b≠0, other-<br>wise leave as is                                                              | USE = XMIS IF T.      |
| Random<br>Number<br>Generation       | RNDU(i) | uniform (0,1) with<br>starting value i.                                                           | X(2) = RNDU(8693).    |
|                                      | RNDG(i) | normal (0,1) with<br>starting value i.                                                            | X(5) = RNDG(8896).    |

-----  
 BMDP STATISTICAL PROGRAMS-----

BMDP1D      SIMPLE DATA DESCRIPTION

Computes summary statistics, lists all data or lists selected cases and stores data on a Save File. The program allows the user to create new grouping variables which are added to the data as additional variables if the data are saved on a Save File.

OUTPUT

- For each variable; mean, std. dev., std error of mean, coef. of variation, largest and smallest values, largest and smallest standardized (z) -scores, range, and total number of acceptable cases.
- For each grouping variable generated; mean, std. dev. and total # of cases grouped.

# BMDP2D - FREQUENCY COUNTS

## CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                                | PARAMETERS       | COMMENT |                                                                |
|------------------------------------------------------------------------------------------|------------------|---------|----------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/TRANSF<br>/SAVE<br>/END<br>/PRINT<br>(Optional) | } see MAIN CHART |         |                                                                |
|                                                                                          |                  | MISS.   | Print cases containing missing values (no/prev.)               |
|                                                                                          |                  | MAX.    | Print cases with values exceeding maximum specified (no/prev.) |
|                                                                                          |                  | MIN.    | Print cases with values less than minimum (no/prev.)           |
|                                                                                          |                  | DATA.   | Print all cases (no/prev.)                                     |

## BMDP2D FREQUENCY COUNT ROUTINE

Counts and lists all distinct values in ascending order and computes summary statistics of selected variables. Values may be truncated or rounded by different amounts for different variables. All computations are performed on the truncated or rounded values.

### OUTPUT

- For each selected variable, # of cases; max, min, range, 1/2 interquartile range, mean, median, mode, std. dev., std. error of the mean, skewness and kurtosis.
- Each distinct value, its frequency, % of cumulative.
- A histogram for each selected variable.
- A line plot of the computed statistics.

## CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                      | PARAMETERS       | COMMENT           |                                                                               |
|--------------------------------------------------------------------------------|------------------|-------------------|-------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/END<br>/COUNT<br>(optional) | } see MAIN CHART |                   |                                                                               |
|                                                                                |                  | TRUNC=(#)#.       | truncate or round the specified variable(s) to the nearest unit given. (none) |
|                                                                                |                  | or<br>ROUND=(#)#. | truncates variable 3 to next lower 10.                                        |
|                                                                                |                  | TRUNC=(3)10.      | (13.82 becomes 14)                                                            |
|                                                                                |                  | ROUND=(3)1.       | Requests a series of location statistics. (no)                                |

## BMDP3D t-TEST AND T SQUARED ROUTINE

Computes one and two sample t-tests along with their associated probability levels. Paired t-test is done by creating a transformed difference variable and performing one sample t-test on this difference variable.

### OUTPUT

- means, std. dev., std. error of means, max. and mins.
- a 20-interval histogram for each group.



- F value for comparison of variances, and pooled separate t-values.
- two-tailed x values for each t and F.
- optionally Mahalanobis D , Hotelling's T squared significance.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                         | PARAMETERS                      | COMMENT                                                              |
|-------------------------------------------------------------------|---------------------------------|----------------------------------------------------------------------|
| <pre> /PROB /INPUT /VARIAB /GROUP /TRANSF /SAVE /END /TEST </pre> | <pre> } see MAIN CHART </pre>   |                                                                      |
|                                                                   | CORR.                           | Prints the within group correlation matrices if desired. (no/prev.)  |
|                                                                   | TITLE='text'.<br>VARIAB = list. | test title if desired.(blank) variables to be tested.(all variables) |
|                                                                   | HOTEL.                          | calculate Hotelling's T squared (no/prev.)                           |
|                                                                   | COMP.                           | Use only complete cases. (no/prev.)                                  |
|                                                                   | GROUP=List.                     | specify specific groups to be compared if not all are desired. (all) |

BMDP4D ALPHANUMERIC FREQUENCY COUNT ROUTINE

Counts the characters found in single column fields on the input data. The program assumes that a single column field is read for each variable in A1 format. All IBM 026 and 029 keypunch characters are considered legal and the frequency of each is counted separately. Strictly a data screening program.

OUTPUT

Frequency counts of all distinct characters.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                | PARAMETERS                                                       | COMMENTS                                                                    |
|----------------------------------------------------------|------------------------------------------------------------------|-----------------------------------------------------------------------------|
| <pre> /PROB /INPUT /VARIAB /END /PRINT (optional) </pre> | <pre> } See MAIN CHART </pre>                                    |                                                                             |
|                                                          | .....No LABEL, CONTENT, CODE, TYPE.<br>.....NAME and LABEL only. |                                                                             |
|                                                          | ALPHA =<br>'Character'.                                          | A character to replace all alphabetic characters when printing data. (none) |
|                                                          | NUMERIC =<br>'character'.                                        | A character to replace all numerical characters. (none)                     |
|                                                          | BLANK =<br>'character'.                                          | A character replacing all blanks. (none)                                    |
|                                                          | PERIOD =<br>'character'.                                         | A character replacing all periods or decimals. (none)                       |
|                                                          | SIGN =<br>'character'.                                           | A character replacing all plus or minus signs. (none)                       |
|                                                          | SYMB =<br>'character'.                                           | A character replacing all symbols. (none)                                   |

## BMDP5D - UNIVARIATE PLOTTING

ILL =                    A character replacing all  
'character'.            illegal values in the data.

PRINT SYMBOL IS '\$'.

---

### BMDP5D            UNIVARIATE PLOTTING

Prints univariate plots. For each plot, cases belonging to one or more groups can be used. Cases from different groups can be identified by distinct letters. Plots can be printed for all groups in one plot or for each group individually. The size of the plot can be controlled.

#### OUTPUT

- A histogram and/or a cumulative histogram.
- A normal probability plot, half-normal probability plot, and/or cumulative distribution plot.
- The total frequency count, mean and standard deviation (for each group in a plot).

#### CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                 | PARAMETERS       | COMMENTS                                                                                                                                                                                    |
|---------------------------------------------------------------------------|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/TRANSF<br>/SAVE<br>/END<br>/PLOT | } see MAIN CHART |                                                                                                                                                                                             |
| (required,<br>may be<br>repeated)                                         |                  |                                                                                                                                                                                             |
|                                                                           | VARIAB = list.   | variables to be plotted.<br>(ALL/prev.)                                                                                                                                                     |
|                                                                           | TYPE = name.     | One or more of: HIST, CHIST,<br>NORM, HALFNORM, DNORM, CUM(HIST<br>/prev.)                                                                                                                  |
|                                                                           | GROUP = list.    | Groups used in one plot if not<br>all. (ALL)                                                                                                                                                |
|                                                                           | SIZE = #1,#2.    | Length of horizontal & vertical<br>axis in characters & lines for<br>NORM, DNORM, HALFNORM, AND CUM.<br>(80,50/prev.)                                                                       |
|                                                                           | SCALE = #1,#2.   | Frequency scale of HIST or CHIST<br>SCALE = 10,2 start printing<br>after a frequency of 10 is<br>achieved and print one symbol<br>every time 2 or more data points<br>are in the cell.(0,1) |
|                                                                           | MAX = #list. .   | Upper limits for the variables,<br>one per variable (none/prev.)                                                                                                                            |
|                                                                           | MIN = #list.     | Lower limits. (none/prev.)                                                                                                                                                                  |

---

### BMDP6D            BIVARIATE PLOTTING

Prints bivariate scatter plots. Cases belonging to one or more groups can be used for each plot. The plots may have frequencies plotted for each point, or distinct letters may be used to indicate group membership. Plots can be printed for all groups in one plot for each group in separate plots.

#### OUTPUT

- Scatter plots of each specified pair of variables using cases of desired groups.
- Mean, std. dev., equation of simple linear regression of each variable on the other.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                                     | PARAMETERS                     | COMMENTS                                                                                                 |
|-----------------------------------------------------------------------------------------------|--------------------------------|----------------------------------------------------------------------------------------------------------|
| <pre> /PROB /INPUT /VARIAB /GROUP /TRANSF /SAVE /END /PLOT (required, can be repeated) </pre> | see MAIN CHART                 |                                                                                                          |
|                                                                                               | XVAR = List.                   | Variables for X-axis<br>(1st var/prev.)                                                                  |
|                                                                                               | YVAR = List.                   | Variables for Y-axis<br>(1st var./prev.)                                                                 |
|                                                                                               | PAIR. or<br>CROSS.             | PAIR tabulates n XVAR<br>by m YVAR. Cross tabulates<br>each XVAR by XVARs (PAIR).                        |
|                                                                                               | GROUP = List.                  | Groups to use in plots if all<br>are not wanted.(all in one plot)                                        |
|                                                                                               | FREQ.<br>CHAR = GROUP.<br>VAR. | FREQ causes frequency plots.<br>GROUP and VAR causes group or<br>variable membership to be<br>plotted.   |
|                                                                                               | STAT.                          | Requests statistics to be<br>printed. (NO STAT/prev.)                                                    |
|                                                                                               | SIZE = #list.                  | Length of the X & Y axis in<br>characters and lines.<br>(70,42/prev.)                                    |
|                                                                                               | MIN = #list.                   | Lower limits, one per var.<br>(none/prev.)                                                               |
|                                                                                               | MAX = #list.                   | Upper limits, one per var.<br>(none/prev.)                                                               |
|                                                                                               | XCOM.                          | All XVARs are plotted on<br>common X-axis.(NO)                                                           |
|                                                                                               | YCOM.                          | All YVARs are plotted on<br>common Y-axis.(NO)                                                           |
|                                                                                               | COM.                           | All plots are plotted in a<br>common frame.(NO)                                                          |
|                                                                                               | SYMB = 'C1','C2'...            | Symbols used for group or<br>variable identification. (1,2,...<br>if CHAR = FREQ, A, B,...<br>otherwise) |

BMDP7D DESCRIPTION OF STRATA WITH HISTOGRAMS AND ANALYSIS OF VARIANCE

Groups the data into a specified number ( 10) of groups based on the values of a grouping variable or on the order of entry of the data. For each variable the program prints histograms for each group, side by side. The number of intervals in the histograms may be specified or the number intervals can be computed by the program. A one- or two-way ANOVA is reported.

# BMDP8D - MISSING VALUE CORRELATION

## OUTPUT

- Histogram for each variable and group.
- Mean, std. dev., count, max, min. for each variable and group.
- One-way or two-way ANOVA for each variable.
- Optional data listing, winsorized mean & confidence intervals by group and correlations.

## CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                                | PARAMETERS          | COMMENTS                                                                   |
|------------------------------------------------------------------------------------------|---------------------|----------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/TRANSF<br>/SAVE<br>/END<br>/PRINT<br>(optional) | }<br>see MAIN CHART |                                                                            |
|                                                                                          | CORR.               | Print correlations of histograms for each group & all groups. (no/prev.)   |
|                                                                                          | DATA.               | Print input data matrix. (no/prev.)                                        |
|                                                                                          | ORDER.              | Print input data in ascending order on the histogram variables. (no/prev.) |
|                                                                                          | WINSOR.             | Print winsorized means. (no/prev.)                                         |
| /HIST<br>(required)                                                                      | GROUP = List.       | Variables stratifying data into groups. (none)                             |
|                                                                                          | VARIAB = List.      | Variables for histograms and ANOVA. (all/prev.)                            |
|                                                                                          | INCR = Values.      | Interval width in histograms, one for each variable. (max-min/INTER).      |
|                                                                                          | INTER = Values.     | Number of intervals in histogram, one for each variable. (0/prev.)         |

## BMDP8D MISSING VALUE CORRELATION

Computes correlation four different ways from data which contains missing values.

- ALL VALUE; the mean of each variable is computed from all existing values of that variable; then deviations from these means are used.
- COVPAIR; each element of the covariance matrix is computed from the existing pairs of values involved, which are then used to compute the correlations.
- CORPAIR; each element of the correlation matrix is computed from existing pairs of values involved.
- COMPLETE; cases with any missing values are not used.

## OUTPUT

- Variance and mean of each variable.

- Optional frequency table of existing pairs, covariance and correlation matrices, sum of weights, matrix of means and variances.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                      | PARAMETERS                                                                                  | COMMENTS                                                                |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/END<br>/PRINT<br>(optional) | see MAIN CHART<br>CONTENT =<br>List.<br>COVA.<br>CORR.<br>FREQ.<br>MEAN.<br>VAR.<br>SUMWTS. | DATA, COVA or CORR.(DATA)                                               |
|                                                                                |                                                                                             | (no/prev.)                                                              |
|                                                                                |                                                                                             | Print covariance matrix.                                                |
|                                                                                |                                                                                             | (YES/prev.)                                                             |
|                                                                                |                                                                                             | Print correlation matrix.                                               |
|                                                                                |                                                                                             | (YES/prev.)                                                             |
|                                                                                |                                                                                             | Print frequency table.                                                  |
|                                                                                |                                                                                             | (YES/prev.)                                                             |
| Print matrix of means.                                                         |                                                                                             |                                                                         |
| (NO/prev.)                                                                     |                                                                                             |                                                                         |
| Print matrix of variances.                                                     |                                                                                             |                                                                         |
| (NO/prev.)                                                                     |                                                                                             |                                                                         |
| Print weight matrix.                                                           |                                                                                             |                                                                         |
| (NO/prev.)                                                                     |                                                                                             |                                                                         |
| /CORR<br>(optional)                                                            | TYPE= list.                                                                                 | One or more of ALLVALUE, COVPAIR, CORPAIR, or COMPLETE. (CORPAIR/prev.) |
|                                                                                | SEG.                                                                                        | Matrices too large are processed in segments.(no/prev.)                 |

BMDP9D MULTIDIMENSIONAL DATA DESCRIPTION

Provides cellwise descriptive statistics for cases when the data are classified multidimensionally.

OUTPUT

- For each variable, cell frequencies, cell means and cell std. devs.
- For each variable across all cells, chi-square test of equality of cell frequencies, F test of cell means and Bartlett's test for homogeneity of cell variances.
- Optional plots of cell mean shifts.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                                   | PARAMETERS                                                                                                                                                                            | COMMENTS                                                                         |
|---------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/GROUP<br>/END<br>/TABULATE<br>(required) | see MAIN CHART<br>GROUP = List.<br>(required)<br>VARIAB = List.<br>MARGIN = 'C1', 'C2', ... Literal map indicating marginals at which variables are described. (all groups).<br>PLOT. | Variables used for indexing or grouping.(none/prev.)                             |
|                                                                                             |                                                                                                                                                                                       | Variables for cellwise description. (all variables).                             |
|                                                                                             |                                                                                                                                                                                       | Literal map indicating marginals at which variables are described. (all groups). |
|                                                                                             |                                                                                                                                                                                       | Request to plot means of each cell. (no/prev.)                                   |
|                                                                                             |                                                                                                                                                                                       |                                                                                  |

BMDP1F TWO-WAY CONTINGENCY TABLES

# BMDP2F - CONTINGENCY TABLE ANALYSIS

Forms two-way contingency tables from raw data or from cell frequencies.

## OUTPUT

- Independence of the row and column classifications.
- Fisher's exact probability for 2 x 2 tables.
- Measures of association and correlation between the classifications as two sets of ranks.
- Optional original frequency, % total table frequency, % of row total, % of column total, expected value, etc.

## CONTROL LANGUAGE USAGE

| PARAGRAPH                            | PARAMETERS                | COMMENTS                                                            |
|--------------------------------------|---------------------------|---------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/CATEG | } see MAIN CHART<br>..... | Required if any ROW, COL, or COND variable has more than 10 values. |
| /TRANSF<br>/SAVE<br>/END             |                           |                                                                     |
| /PRINT<br>(optional)                 | OBS.                      | Print observed frequency table. (YES/prev.)                         |
|                                      | EXP.                      | Print expected cell values. (NO/prev.)                              |
|                                      | ROWP.                     | Print row percentages. (NO/prev.)                                   |
|                                      | COLP.                     | Print column percentages. (NO/prev.)                                |
|                                      | TOLP.                     | Print percentages of total frequency. (NO/prev.)                    |
|                                      | DIF.                      | Print differences. (NO/prev.)                                       |
|                                      | STAND.                    | Print standardized deviations. (NO/prev.)                           |
|                                      | ADJ.                      | Print adjusted standardized deviations. (NO/prev.)                  |
|                                      | SMOOTH.                   | Print smoothed values. (NO/prev.)                                   |
|                                      | FREEMAN.                  | Print table of Freeman-tukey deviates. (NO/prev.)                   |
| /STAT<br>(optional)                  | CHISQ.                    | Print Chi-square and Yate's correlation. (yes/prev.)                |
|                                      | CONT.                     | Print C, phi, Cramers V, alfa, etc. (no/prev.)                      |
|                                      | LARCHI.                   | Print likelihood ratio test. (no/prev.)                             |
|                                      | FISH.                     | Print Fisher's exact prob. (no/prev.)                               |
|                                      | TETRA.                    | Print tetrachoric correlation. (no/prev.)                           |
|                                      | CORR.                     | Print product-moment correlation. (no/prev.)                        |
|                                      | SPEAR.                    | Print Spearman's corr. coef. (no/prev.)                             |
|                                      | GAM.                      | Print gamma, Kendall's tau, and Stuart's tau Somer's D. (no/prev.)  |
|                                      | LAM.                      | Print lambda, and lambda asym. (no/prev.)                           |
|                                      | LSTAR.                    | Print Goodman and Kuskal lambda asym. (no/prev.)                    |
|                                      | TAU.                      | Print Goodman and Kuskal tau asym. (no/prev.)                       |
|                                      | UNC.                      | Print measures of uncertainty U and Uasym. (no/prev.)               |
|                                      | MCN.                      | Print McNemar's test of symmetry (no/prev.)                         |

|                            |                 |                                                                                                  |
|----------------------------|-----------------|--------------------------------------------------------------------------------------------------|
| /TABLE                     | COL=list.       | Identification of col. variables (none)                                                          |
|                            | ROW=list.       | Identification of row variables. (none)                                                          |
| (Required may be repeated) | PAIR. OR CROSS. | PAIR tabulates variable pairs in order. CROSS tabulates each row var. vs. each col. var. (PAIR). |
|                            | COND=list.      | Names or numbers of table conditioning variables. (none)                                         |
|                            | COUNT=list.     | Variables used to sum freq. counts in ROW, COL and/or COND tables. (none)                        |
|                            | MIN=#.          | Minimum expected frequency for collapsing tables. (0)                                            |
|                            | DELTA=#.        | Value added to each cell frequency before calculating statistics. (0)                            |

-----

BMDP2F TWO-WAY CONTINGENCY TABLES -- EMPTY CELLS AND THE IDENTIFICATION OF DEPARTURES FROM INDEPENDENCE.

Analyzes two-way frequency tables. Differs from BMDP1F in that:

- Cells may be excluded from the table by declaring them empty. This allows a test of independence using all cells not declared to be empty and is identical to the usual Chi-square when there are no empty cells.
- Cells or patterns of cells that contribute to a departure from independence or quasi-independence may be identified.

This program does not compute measures of association, correlation, or prediction because of empty cells.

OUTPUT

- Observed frequency table and associated Chi-square.
- Expected values for each cell.
- Stepwise chi-square, cell elimination criteria, previously eliminated cells.
- Final summary table.
- Optional tables.

CONTROL LANGUAGE USAGE

| PARAGRAPH                   | PARAMETER              | COMMENTS                                                                                                   |
|-----------------------------|------------------------|------------------------------------------------------------------------------------------------------------|
| -----                       |                        |                                                                                                            |
| /PROB                       | } see MAIN CHART       |                                                                                                            |
| /INPUT                      |                        |                                                                                                            |
| /VARIAB                     |                        |                                                                                                            |
| /CATEG                      |                        | ..... Required if any ROW, COL or COND variable has more than 10 values.                                   |
| /TRANSF                     |                        | CONTENT=name. One only of DATA or OBS.(DATA)                                                               |
| /SAVE                       |                        |                                                                                                            |
| /END                        |                        |                                                                                                            |
| /TABLE                      | ROW,COL,PAIR or CROSS. | see BMDP1F                                                                                                 |
| (Required, may be repeated) | COND,COUNT,OBS,DELTA   |                                                                                                            |
|                             | FIT.                   | Print table of fitted cell values.(no)                                                                     |
|                             | EMPTY (i)= list.       | Row and column cell in indices empty a priori. (i) refers to the table number. EMPTY (3) = 1,1,2, 3.(none) |
|                             | CONV=#.                | Convergence criterion to find fitted value. (.05).                                                         |
|                             | ITER=#.                | Max iterations for fitting quasi-independent model. (20)                                                   |

# BMDP3F - MULTIWAY CONTINGENCY TABLES

|                                              |                          |                                                                                                                                                                                                                                                                                                                                |
|----------------------------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| /STEP<br>(optional,<br>may be re-<br>peated) | CRIT=name.<br>(required) | Criterion for cell elimination<br>CHISQ chi-square, LRCHI likelihood<br>ratio, USTAN largest absolute<br>std. quasi-residual, QUADJ<br>largest absolute adjusted std.<br>quasi-residual, STAN max. abso-<br>lute adjusted std. residual, ADJ<br>max. absolute adjusted std.<br>residual, DIF max. absolute<br>residual. (none) |
|                                              | MAXST=#.                 | Maximum # of steps.(degrees of<br>freedom)                                                                                                                                                                                                                                                                                     |
|                                              | ALPHA=#.                 | Level of significance for testing<br>quasi-independence.(.05)                                                                                                                                                                                                                                                                  |
|                                              | ELIM=#.                  | Criterion limit to eliminate<br>cells.(1.0)                                                                                                                                                                                                                                                                                    |
|                                              | REEN=#.                  | Criterion limit to reenter cells.<br>(1.0)                                                                                                                                                                                                                                                                                     |
|                                              | PRINT=#.                 | Steps at which tables below are<br>printed.(none)                                                                                                                                                                                                                                                                              |
|                                              | FIT.                     | Print table of fitted values.(no)                                                                                                                                                                                                                                                                                              |
|                                              | CHISQ.                   | Print resulting chi-square when<br>cell is eliminated.(no)                                                                                                                                                                                                                                                                     |
|                                              | LRCHI.                   | Print table of likelihood ratio<br>chi-square.(no)                                                                                                                                                                                                                                                                             |
|                                              | QUSTAN.                  | Print table of quasi-std. devi-<br>ations.(no)                                                                                                                                                                                                                                                                                 |
|                                              | QUADJ.                   | Print table of quasi-adjusted<br>std. devs.(no)                                                                                                                                                                                                                                                                                |
|                                              | STAN.                    | Print table of standard residuals<br>(no)                                                                                                                                                                                                                                                                                      |
|                                              | ADJ.                     | Print table of adjusted std.<br>residuals.(no)                                                                                                                                                                                                                                                                                 |
|                                              | DIF.                     | Print table of residuals.(no)                                                                                                                                                                                                                                                                                                  |

## BMDP3F MULTIWAY CONTINGENCY TABLES

Forms and analyzes multi-way contingency tables. The purpose of the analysis is to obtain a description of the relationships between factors of the table, either by forming a model for the data or by testing and ordering the importance of the interactions between the factors. The analysis is based on fitting a log-linear model to the cell frequencies.

Input may be the usual variables-by-case format, cell indices and frequency counts, or a multiway frequency table.

### OUTPUT

- The original frequency table and marginal totals.
- A test for each k that the k-factor interactions are zero.
- An ANOVA-like table of likelihood ratio Chi-square stats.
- Estimated parameters and tests of fit for a user defined model.

### CONTROL LANGUAGE USAGE

| PARAGRAPH                                       | PARAMETER        | COMMENTS                                                                       |
|-------------------------------------------------|------------------|--------------------------------------------------------------------------------|
| -----                                           |                  |                                                                                |
| /PROB<br>/INPUT<br>/VARIAB<br>/CATEG<br>/TRANSF | } see MAIN CHART |                                                                                |
| /SAVE                                           | CONTENT=name     | Required if any ROW, COL or COND<br>var. has more than 10 different<br>values. |
| /END<br>/INPUT                                  | TABLE=#list.     | One only of DATA or OBS.(DATA)<br>Number of categories for each                |



|                                             |                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|---------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| /TABLE<br>(Required,<br>may be<br>repeated) | INDEX=List.<br>SYMB=List.<br>COUNT=List.<br>COND=List.<br>OBS.<br>MARG=#.<br>DELTA=Value<br>INTER=#.<br>CONV=#.<br>SIMUL.<br>ASSOC.<br>ALL.<br>MODEL=List.<br>EXP.<br>STAN.<br>FREEMAN.<br>LAMB.<br>BETA.<br>ORTH=List.<br>PROB=#.<br>ADD.<br>DELETE. | index if a frequency table is<br>input fastest index first.(none)<br>Variables used as indices for<br>table.(none)<br>Symbols assigned to INDEX vari-<br>ables, in output.(none)<br>Variables that are frequency<br>counts.(none)<br>For every level of these vari-<br>ables, a table is formed.(none)<br>Print observed frequency table.<br>(yes/prev.)<br>Print marginal subtables up to<br>order k.(0/prev.)<br>Add Delta to each frequency be-<br>fore analysis.(0/prev.)<br>Max # of iterations.(20/prev.)<br>Criterion for convergence of<br>model fitting.(.1/prev.)<br>Calculate tests-of-fit of models<br>of full order.(yes/prev.)<br>Calculate tests of partial and<br>marginal association.(yes/prev.)<br>Calculate all models if 2 or 3-<br>way table.(no/prev.)<br>Hierarchical model to be fitted.<br>(none)<br>Print table of fitted (expected)<br>values.(no)<br>Print std. residuals.(no)<br>Print Freeman-Tukey deviates.(no)<br>Est. log-linear parameters of<br><br>Est. multiplicative parameters of<br>MODEL.(no)<br>Fit orthogonal polynomials for<br>SYMBOLS specified.(none)<br>Print LAMB, BETA, ORTH only if<br>test-of-fit of model is greater<br>than this value. (.05)<br>Form models that contain MODEL as<br>a subser.(no)<br>Form models contained by MODEL.<br>(no) |
| /FIT<br>(Optional,<br>may be<br>repeated)   | MODEL=List.<br>EXP.<br>STAN.<br>FREEMAN.<br>LAMB.<br>BETA.<br>ORTH=List.<br>PROB=#.<br>ADD.<br>DELETE.                                                                                                                                                | Hierarchical model to be fitted.<br>(none)<br>Print table of fitted (expected)<br>values.(no)<br>Print std. residuals.(no)<br>Print Freeman-Tukey deviates.(no)<br>Est. log-linear parameters of<br><br>Est. multiplicative parameters of<br>MODEL.(no)<br>Fit orthogonal polynomials for<br>SYMBOLS specified.(none)<br>Print LAMB, BETA, ORTH only if<br>test-of-fit of model is greater<br>than this value. (.05)<br>Form models that contain MODEL as<br>a subser.(no)<br>Form models contained by MODEL.<br>(no)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

-----

BMDP1M      CLUSTER ANALYSIS ON VARIABLES

Clusters similar variables by using an initial measure of association between pairs of variables to form a cluster of the two most similar variables and then using an amalgamation rule to form further clusters.

OUTPUT

- Summary table for the clustering process.
- Tree diagram of clusters.
- Table of similarity or distance matrix scaling.
- Explanation of tree diagram.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                              | PARAMETER                        | COMMENTS                                                       |
|--------------------------------------------------------|----------------------------------|----------------------------------------------------------------|
| -----                                                  |                                  |                                                                |
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/END | }<br>TYPE=name<br>see MAIN CHART | One only of DATA, COVA, CORR, ANG,<br>SIMI, DIST. (DATA/prev.) |

## BMDP2M - CLUSTER ANALYSIS ON CASES

|                      |                  |                                                                                                                                             |
|----------------------|------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| /PRINT<br>(Optional) | SHADE.           | Print shaded correlation matrix.<br>(no/prev.)                                                                                              |
|                      | CORR.            | Print correlation matrix.<br>(no/prev.)                                                                                                     |
| /PROC<br>(Optional)  | VARIAB=<br>List. | Variables to be clustered.(all in<br>use/prev.)                                                                                             |
|                      | MEAS=name.       | Association measure: CORR<br>Correlation, ABCORR absolute<br>correlation, ANG arccos (Corre-<br>lation), ABSANG arccos (abs corr)<br>(corr) |
|                      | LINK=name.       | Amalgamation rule, one only:<br>SINGLE, COMP, AVE.(SINGLE/prev.)                                                                            |

### BMDP2M CLUSTER ANALYSIS ON CASES

Clusters cases on the basis of one of four available distance measures.

#### OUTPUT

- A diagram indicating clustering of cases.
- Optional standardized input data, initial distance case matrix and vertical clustering diagram.

#### CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                     | PARAMETER        | COMMENTS                                                                                        |
|-------------------------------------------------------------------------------|------------------|-------------------------------------------------------------------------------------------------|
| -----                                                                         |                  |                                                                                                 |
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/END<br>/PROC<br>(Optional) | } see MAIN CHART |                                                                                                 |
|                                                                               | SUMOFSQ          | Square root of sums of squares<br>of differences of distances be-<br>tween clusters.(yes/prev.) |
|                                                                               | SUMOFP=p.        | pth root of sum of pth powers of<br>diff. as distance measure.(2)                               |
|                                                                               | CHISQ.           | Chi-square as distance.(no)                                                                     |
|                                                                               | PHISQ.           | Chi-square/n as distance.(no)                                                                   |
|                                                                               | STAND.           | Standardize the input data.(yes)                                                                |
|                                                                               | WEIGHT=name.     | Variable containing case weights<br>(unit weights). (no weights/prev.)                          |
| /PRINT<br>(Optional)                                                          | DATA.            | Print std. data. (no/prev.)                                                                     |
|                                                                               | DIST.            | Print initial std. distances.<br>(no/prev.)                                                     |
|                                                                               | REG. or VERT.    | Print horizontal or vertical<br>triangle tree diagram.(VERT/<br>prev.)                          |
|                                                                               | AMALG.           | Print table of distances, case<br>weights and averages by step<br>(yes/prev.)                   |
|                                                                               | SHADE.           | Print shaded form of distance<br>matrix.(no/prev.)                                              |

### BMDP3M BLOCK CLUSTERING

Simultaneously clusters cases and variables of a data matrix. Three types of clusters are thus discovered: clusters of data values (the blocks), clusters of cases, and clusters of variables.

#### CONTROL LANGUAGE USAGE

| PARAGRAPH       | PARAMETER | COMMENTS |
|-----------------|-----------|----------|
| -----           |           |          |
| /PROB<br>/INPUT | }         |          |

```

/VARIAB INTER=#list. Number of intervals, one per
/TRANSF } variable (first 35 values)
/SAVE }
/END } see MAIN CHART
/PASS } TOTAL=#. Number of passes on rows and
(Optional) } THRESH=#list. Number indicating threshold at
 each pass. (increase from 2/
 TOTAL to 1)

```

BMDP4M FACTOR ANALYSIS

Performs a factor analysis of either a correlation matrix or a covariance matrix. The input may be raw data, correlation matrix, covariance matrix, factor loadings or factor score coefficients.

OUTPUT

- Means, std. dev., coefficients of variation, maximums and minimums, standard scores.
- Correlation matrix, covariance matrix.
- Factor loadings and their plots.
- Squared multiple correlations and Kaiser's measure of sampling adequacy.
- Factor correlations and factor structure.
- Factor scores, factor score coefficients, factor score covariances and factor score plots.

CONTROL LANGUAGE USAGE

| PARAGRAPH                   | PARAMETER     | COMMENTS                                                    |
|-----------------------------|---------------|-------------------------------------------------------------|
| /PROB                       |               |                                                             |
| /INPUT                      |               |                                                             |
| /VARIAB                     | SCORE=list.   | Variables used for factor scores. (none)                    |
| /TRANSF                     | WEIGHT=var.   | Variable containing case weight. (no weight/prev.)          |
| /SAVE                       |               |                                                             |
| /END                        |               |                                                             |
| /INPUT                      | TYPE=name.    | One of DATA, CORR, COVA., LOAD, FSCF.(DATA or CONTENT=)     |
| (Required on first problem) | CONTENT=name. | One of DATA, CORR, COVA, LOAD, FSCF. (DATA)                 |
|                             | FACTOR=#.     | Number of factors when TYPE=LOAD or FSCF. (NONE)            |
| /SAVE                       | SHAPE=name.   | One of SQUARE, LOWER. (SQUARE)                              |
| (Optional)                  | CONTENT=name. | One or more of DATA, COVA, UFLD, RFLD, FCOR. (DATA)         |
| /PRINT                      | CASE=#.       | Number of cases for which data is printed. (5/prev.)        |
| (Optional)                  | FSCORE=#.     | Number of factors for printing factor scores. (all)         |
|                             | HILEV=#.      | High level factor loading. (.5/prev.)                       |
|                             | LOLEV=#.      | Low level factor loading. (.25/prev.)                       |
|                             | STAN.         | Print standard scores.(no/prev.)                            |
|                             | COVA.         | Print covariance matrix. (no/prev.)                         |
|                             | CORR.         | Print correlation matrix. (yes/prev.)                       |
|                             | INV.          | Print inverses.(no/prev.)                                   |
|                             | PART.         | Print partial correlations. (no/prev.)                      |
|                             | FSTR.         | Print factor structure matrix (oblique rotation).(no/prev.) |

# BMDP6M - CANONICAL CORRELATION

|                       |              |                                                                                                                                                                                                                                     |
|-----------------------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                       | SHADE.       | Print shaded correlation.                                                                                                                                                                                                           |
|                       | RESI.        | Print residual correlations.<br>(no/prev.)                                                                                                                                                                                          |
| /PLOT<br>(Optional)   | INITIAL=#.   | Number of unrotated factors for<br>factor loadings plots.(4)                                                                                                                                                                        |
|                       | FINAL=#.     | Number of rotated factors (4)                                                                                                                                                                                                       |
|                       | FSCORE=#.    | Number of factors for factor<br>score plots. (3)                                                                                                                                                                                    |
| /FACTOR<br>(Optional) | CONST=value. | Value limiting factors to<br>eigenvalues >CONST.(1/prev.)                                                                                                                                                                           |
|                       | NUMBER=#.    | MAXIMUM # OF FACTORS OBTAINED.<br>(# of variables)                                                                                                                                                                                  |
|                       | METHOD=name. | One only of PCA principle compo-<br>nents, MLFA maximum likelihood,<br>LJIFFY Kaiser's second generation<br>Little Jiffy, PFA principle factor<br>analysis.(PCA).                                                                   |
|                       | ITERATE=#.   | Maximum # of iterations.(25/prev.)                                                                                                                                                                                                  |
|                       | EPS=#.       | Convergence criteria.(.001/prev.)                                                                                                                                                                                                   |
|                       | COMMUN=name. | One of the following communality<br>options: UNALT unaltered, SMCS<br>squared multiple correlations,<br>MAXROW, maximum absolute row<br>values of the correlation matrix.<br>(UNALT for PCA, SMCS otherwise/<br>prev.)              |
|                       | FORM=name.   | One of the following for factor<br>form: CORR factor correlation<br>matrix, COVA covariance matrix,<br>OCORR correlation matrix about<br>the origin, OCOVA covariance<br>matrix about the origin.(CORR/<br>prev.)                   |
| /ROTATE<br>(Optional) | METHOD=name. | NONE no rotation, VMAX varimax,<br>DQUART direct quartimin, ORMAX<br>quartimax, EQMAX equamax, ORTHOG<br>orthogonal with gamma below,<br>DOBLI direct oblimin with gamma,<br>ORTHOB orthoblique. (VMAX except<br>with LJIFFY/prev.) |
|                       | GAMMA=value. | Parameter for METHOD=ORTHOG or<br>DOBLI. (1 for ORTHOG, 0.0 for<br>DOBLI)                                                                                                                                                           |
|                       | CONST=value. | Convergence criteria for rotation<br>(.00001/prev.)                                                                                                                                                                                 |
|                       | MAXIT=#.     | Max. # of iterations for rotation.<br>(50/prev.)                                                                                                                                                                                    |
|                       | NORMAL.      | Request Kaiser's normalization.<br>(yes/prev.)                                                                                                                                                                                      |

## BMDP6M CANONICAL CORRELATION ANALYSIS

Computes canonical correlation analysis for two sets of variables. Input can be data, a covariance or a correlation matrix. Computations can be done in single or double precision.

### OUTPUT

- Means, std. devs., coefficients of variation, skewness kurtosis, mins and maxs.
- Correlations, covariances and canonical correlations.
- Eigenvalues associated with each pair of canonical variables.
- Coefficients of canonical variables and canonical variables.
- Correlation of variables with canonical variable and plots

### CONTROL LANGUAGE USAGE

| PARAGRAPH                                        | PARAMETER                       | COMMENTS                                                                                                                                                                  |
|--------------------------------------------------|---------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB                       | SHAPE=name.                     | SQUARE, LOWER. (SQUARE)                                                                                                                                                   |
|                                                  | TYPE=name.                      | One of: DATA,COVA,CORR.(DATA)                                                                                                                                             |
| /TRANSF<br>/SAVE<br>/END<br>/PRINT<br>(Optional) | WEIGHT=V.                       | Variable containing case weight.<br>(none/prev.)                                                                                                                          |
|                                                  | see MAIN CHART<br>CONTENT=list. | One or both DATA,COVA. (DATA)                                                                                                                                             |
| /PLOT<br>(Optional)                              | CASE=#.                         | Number of cases to print data.<br>(5/prev.)                                                                                                                               |
|                                                  | MATR=list.                      | CORR correlations, COVA covari-<br>ances, CANV canonical variables,<br>COEF canonical var. regression<br>coefficients, LOAD canonical vari-<br>able loadings. (CORR,LOAD) |
| /CANON<br>(Required)                             | XVAR=names.                     | Names or subscripts of x-axis<br>vars.(none)                                                                                                                              |
|                                                  | YVAR=names.<br>SIZE=#1,#2.      | Y-axis variables. (none)<br>#1 is number of columns. #2 is<br>number of rows.(50,50/prev.)                                                                                |
|                                                  | FIRST=names.                    | First set of variables.(none)<br>(required)                                                                                                                               |
|                                                  | SECOND=names.                   | Second set of variables. (none)<br>(required)                                                                                                                             |
|                                                  | TOL=#.                          | Tolerance of matrix inversion.<br>(.01/prev.)                                                                                                                             |
|                                                  | PREC=name.                      | SINGLE single precision, DOUBLE<br>double precision.(DOUBLE/prev.)                                                                                                        |
|                                                  | ZERO.                           | Assume means zero. (NO ZERO/prev.)                                                                                                                                        |
|                                                  | CONST=#.                        | Limit that canonical correlation<br>must be greater than.(0.0/prev.)                                                                                                      |
|                                                  | NUMBER=#.                       | Max. # of canonical variables.<br>(# in smaller of two variable sets<br>FIRST and SECOND)                                                                                 |

#### BMDP7M STEPWISE DISCRIMINANT ANALYSIS

Performs stepwise multiple group discriminant analysis.

#### OUTPUT

- Group means and std. devs.
- F statistics, Wilk lambda and Magalanobis' D square.
- Classification functions, classification matrices and % correct classification.
- Coefficients for canonical discriminant functions, canonical correlations and canonical variables.
- Plots of the first two canonical variables.

#### CONTROL LANGUAGE USAGE

| PARAGRAPH                            | PARAMETER                  | COMMENTS                                                                         |
|--------------------------------------|----------------------------|----------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP | see MAIN CHART<br>GROUP=#. | Number of groups if >10 and no<br>GROUP paragraph.(none/prev.)                   |
|                                      | USE=list.                  | Groups for within-group cross-<br>product matrix.(all groups/prev.)              |
| /TRANSF                              | PRIOR=#list.               | Prior probabilities used to com-<br>pute posterior probabilities.<br>(1.0/prev.) |
| /SAVE                                | NCAN=#.                    | Number of canonical variables<br>saved.(# of groups - 1/prev.)                   |
| /END<br>/PRINT<br>(Optional)         | STEP.                      | Print detailed stats of each step.<br>(yes/prev.)                                |
|                                      | CLASS=#list.               | Print classification functions                                                   |

BMDPAM - MISSING DATA

|                     |                |                                                                                                                                                                                                      |
|---------------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                     | POST.          | when number of variables entered equals # on this list.(none/prev.)                                                                                                                                  |
|                     | POINT.         | Posterior probabilities and D squared are printed.(yes/prev.)                                                                                                                                        |
|                     | WITH.          | Values of canonical variables are printed.(yes/prev.)                                                                                                                                                |
|                     |                | Print pooled with-group covariance matrix. (no/prev.)                                                                                                                                                |
| /PLOT               | CANON.         | Plot of first two canonical variables. (yes/prev.)                                                                                                                                                   |
|                     | GROUP=list.    | Names or subscripts of groups used for canonical var. plots. (all groups/prev.)                                                                                                                      |
|                     | CONTR.         | Canonical variables are based on specified contrasts.(no/prev.)                                                                                                                                      |
| /DISC<br>(Optional) | TOL=#.         | Number (.001,1) used to guard against singularity.(.01/prev.)                                                                                                                                        |
|                     | ENTER=#1,#2.   | Limiting value of F to enter. (4.4/prev.)                                                                                                                                                            |
|                     | REMOVE=#1,#2.  | Limiting value of F to remove. (3.99,3.99/prev.)                                                                                                                                                     |
|                     | METHOD=1 or 2. | Method of deletion; 1 use F to remove, 2 delete to obtain better prediction.(1/prev.)                                                                                                                |
|                     | LEVEL=#list.   | List of levels for forcing variables into equation. Zero level not entered. Lower levels entered before higher levels. Levels list to correspond to variable indices. (next higher than any in list) |
|                     | FORCE=#.       | Levels up to which variables are to be forced into the equation. (0/prev.)                                                                                                                           |
|                     | STEP=#.        | Max. # of steps allowed.(twice # of variables/prev.)                                                                                                                                                 |
|                     | CONTR=values.  | Values specifying a contrast used to direct stepping. (contrast to equality of group means/prev.)                                                                                                    |
|                     | JACK.          | Compute posterior probabilities. (no/prev.)                                                                                                                                                          |

BMDPAM DESCRIPTION AND ESTIMATION OF MISSING DATA

Describes the pattern of missing values for multivariate data and provides estimates of the missing data. Estimates can be means; regression values on the available variable with which it is most highly correlated; regression on a set of variables highly correlated or regression on all available variables.

OUTPUT

- Pattern of missing data.
- Squared multiple correlations of each variable with all others.
- Missing data estimates.
- Bivariate plots.
- Correlation matrix, eigenvalues, covariance matrix.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                | PARAMETER        | COMMENTS                           |
|----------------------------------------------------------|------------------|------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/TRANSF<br>/SAVE | } see MAIN CHART |                                    |
|                                                          | CONTENT=list.    | One or more of DATA,COVA CORRDICh. |

|            |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|            |              | (DATA/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| /VARIAB    | WEIGHT.      | Variable containing case weights.                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| (Optional) |              | (none)                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|            | GROUP=v.     | Variable used for grouping.(none/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| /PRINT     | CASE=#.      | # of cases to print data.(5/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| (Optional) | MATR=List.   | Matrices to print: CORR correlating, EIGEN eigenvalues, PAT pattern of missing data, FREQ sample size & % missing for each pair of vars. DIS Mahalanobis' distance from each case to center. COVA covariances, SUMWT pairwise sums of weights, INVCV inverse of correlation matrix, printed as a correlation matrix, CORRDICh correlations computed using the dichotomies: 1.0 if the value is acceptable and 0.0 otherwise.(CORR, EIGEN, PAT, EST, FREQ, DIS, CORRDICh/prev.) |
| /PLOT      | XVAR=names.  | X-axis variables.(none)                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| (Optional) | YVAR=names.  | Y-axis variables.(none)                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|            | SIZE=#,#.    | Length of X & Y axis.(50,50/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| /EST       | METHOD=name. | MEAN means, SINGLE regression on correlated variable, TWO-STEP or STEP stepwise regression estimate REGRES regression on all variables NONE no estimates.(none/prev.)                                                                                                                                                                                                                                                                                                          |
| (Optional) |              | F to enter.(4/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|            | ENTER=#.     | Algorithm for computing univariate statistics; COMPLETE complete cases only, ALLVALUE all existing values(ALLVALUE/prev.)                                                                                                                                                                                                                                                                                                                                                      |
|            | TYPE=name.   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|            | VLIM=#.      | Maximum number of missing variables for each case.(variables used -1/prev.)                                                                                                                                                                                                                                                                                                                                                                                                    |
|            | CLIM=%.      | Maximum % of cases missing for any variable.(50/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                         |
|            | TOL=#.       | Tolerance (0,1) for matrix inversion.(.0001/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                             |
|            | ZERO.        | Assume means are zero.(no/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|            | MIN.         | Treat values below minimum limit as missing.(no/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                         |
|            | MAX.         | Treat values above maximum limit as missing(no/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                          |
|            | GROUP.       | Estimate missing grouping variable (no/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                  |

BMDP1R      MULTIPLE LINEAR REGRESSION

Computes a multiple linear regression equation on all data and on groups or subsets of data. Equations with or without an intercept can be chosen.

OUTPUT

- Means, std. dev., mins and max for variables.
- Multiple R and std. error of estimate.
- ANOVA table on regression.
- Optional, covariance and/or correlation matrix. Scatter plots, normal and detrended normal probability plots of residuals. Partial residual plots. Residual, predicted value and data for each case.

CONTROL LANGUAGE USAGE

| PARAGRAPH | PARAMETER | COMMENTS |
|-----------|-----------|----------|
| -----     | -----     | -----    |

# BMDP2R - STEPWISE MULTIPLE LINEAR REGRESSION

|                                                                               |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------------------------------------------------------------------------|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/SAVE<br>/END<br>/PRINT<br>(Optional) | } | see MAIN CHART<br>CONTENT=list.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | TYPE=name.      One of: DATA,COVA,CORR. (DATA)<br>One or both of DATA,COVA. (DATA)<br>COVA.            Print covariance matrix.(no/prev.)<br>CORR.            Print correlation matrix.(no/<br>prev.)<br>DATA.            Print input data, residuals and<br>predicted values. (no/prev.)<br>ANOVA.            Print ANOVA for each var.(yes/<br>prev.)<br>STEP.             Print full output at each step.<br>(yes/prev.)<br>PARTIAL.         Print partial correlations.<br>(yes/prev.)<br>COEF.             Print regression coefficients.<br>(yes/prev.)<br>FRATIO.<br>SUMMARY.        Print F ratios. (no/prev.)<br>RREG.             Print multiple R, R squared, F-to-<br>enter and F-to-remove.(yes/prev.)<br>Print corr. between reg. coef.<br>(no/prev.) |
| /PLOT<br>(Optional)                                                           |   | VARIAB=list.    Variables for plotting.<br>(none/prev.)<br>NORMAL.or<br>DNORMAL.        Print normal or detrended nor-<br>mal plots.(no/prev.)<br>PREP=list.        Variables for partial residual<br>plots.(none/prev.)<br>SIZE=#1,#2.      Length of X & Y axis. (50,50/<br>prev.)<br>RESID.            Plot residuals and squared re-<br>siduals against predicted.<br>(no/prev.)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| /REGRES<br>(Required,<br>may be<br>repeated)                                  |   | TITLE='text'.    Regression title <80 cols.(blank)<br>DEPEND=name.    Dependent variable; required<br>first.(none/prev.)<br>INDEP=list.      Independent variables.(all but<br>dependent/prev.)<br>TOL=#.            Tolerance to guard against singu-<br>larity.(.01/prev.)<br>ENTER=#1,#2.    Limits to F-to-enter.(4,4/prev.)<br>REMOVE=#1,#2.   Limits to F-to-remove.(3.9,3.9/<br>prev.)<br>METHOD=name.      Stepping algorithm F: based on<br>same as F with variable inter-<br>change, R based on increase in<br>multiple R, RSWAP same as R with<br>variable interchange.(F/prev.)<br>TYPE=name.        ZERO zero intercept, NONZERO,<br>FLOAT treat intercept as a vari-<br>able to be entered only if sig-<br>nificant. (NONZERO/prev.)<br>LEVEL=#list.     Number list, corresponding to<br>variable order, specifying level<br>for forcing entry. Zeros will not<br>be entered. Lowest level entered<br>first. (highest + 1)<br>FORCE=#.           Level through which variables will<br>be forced into equation. (none/<br>prev.)<br>STEP=#.           Maximum steps. (twice # of vari-<br>ables/prev.) |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

---

## BMDP2R      STEPWISE MULTIPLE LINEAR REGRESSION

Estimates the parameters of multiple linear regression in a stepwise manner. Variables can be forced into the equation and equations without intercepts can be used.

### OUTPUT

- Covariance and/or correlation matrix.



- Coefficients, std. coef., std. errors, F ratios, partial correlations.
- Summary tables for F-ratios, partial correlations and coefficients.
- Predicted values, residuals and data for each case.
- Scatter plots of predicted and observed values versus selected variables.

#### CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                          | PARAMETER                                                                                                                                                                               | COMMENTS                                                                                                                                               |                                                                                                                                          |
|------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br><br>/TRANSF<br>/SAVE<br>/END<br>/PRINT<br>(Optional) | TYPE=name.<br>WEIGHT=name.<br><br>see MAIN CHART<br>CONTENT=list.                                                                                                                       | Any one of DATA,COVA,CORR(DATA)<br>Variable containing case weights.<br>(none/prev.)                                                                   |                                                                                                                                          |
|                                                                                    |                                                                                                                                                                                         | One or both of DATA,COVA(DATA).                                                                                                                        |                                                                                                                                          |
|                                                                                    |                                                                                                                                                                                         | COVA.<br>CORR.<br>DATA.                                                                                                                                | Print covariance matrix.(no/prev.)<br>Print correlation matrix.(no/prev.)<br>Print input data, residuals and predicted values.(no/prev.) |
|                                                                                    |                                                                                                                                                                                         | ANOVA.                                                                                                                                                 | Print ANOVA for each variable.<br>(yes/prev.)                                                                                            |
|                                                                                    |                                                                                                                                                                                         | STEP.                                                                                                                                                  | Print full output at each step.<br>(yes/prev.)                                                                                           |
|                                                                                    |                                                                                                                                                                                         | PARTIAL.                                                                                                                                               | Print partial correlations.<br>(yes/prev.)                                                                                               |
|                                                                                    |                                                                                                                                                                                         | COEF.                                                                                                                                                  | Print regression coefficients.<br>(yes/prev.)                                                                                            |
|                                                                                    |                                                                                                                                                                                         | FRATIO.<br>SUMMARY.                                                                                                                                    | Print F ratios.(no/prev.)<br>Print multiple R, R squared, F-to-enter and F-to-remove.(yes/prev.)                                         |
|                                                                                    |                                                                                                                                                                                         | RREG.                                                                                                                                                  | Print corr. between reg. coef.<br>(no/prev.)                                                                                             |
|                                                                                    | /PLOT<br>(Optional)                                                                                                                                                                     | VARIAB=list.                                                                                                                                           | Variables for plotting.<br>(none/prev.)                                                                                                  |
|                                                                                    |                                                                                                                                                                                         | NORMAL.or<br>DNORMAL<br>PREP=list.                                                                                                                     | Print normal or detrended normal plots.(no/prev.)<br>Variables for partial residual plots.(none/prev.)                                   |
|                                                                                    |                                                                                                                                                                                         | SIZE=#1,#2.<br>RESID.                                                                                                                                  | Length of X&Y axis.(50,50/prev.)<br>Plot residuals and squared residuals against predicted.<br>(no/prev.)                                |
|                                                                                    |                                                                                                                                                                                         | /REGRES<br>(Required,<br>may be<br>repeated)                                                                                                           | TITLE='text'.<br>DEPEND=name.                                                                                                            |
| INDEP=list.                                                                        |                                                                                                                                                                                         |                                                                                                                                                        | Independent variables.(all but dependent/prev.)                                                                                          |
| TOL=#.                                                                             | Tolerance to guard against singularity.(.01/prev.)                                                                                                                                      |                                                                                                                                                        |                                                                                                                                          |
| ENTER=#1,#2.<br>REMOVE=#1,#2.                                                      | Limits to F-to-enter.(4,4/prev.)<br>Limits to F-to-remove.(3.9,3.9/prev.)                                                                                                               |                                                                                                                                                        |                                                                                                                                          |
| METHOD=name.                                                                       | Stepping algorithm F: based on F-to-remove & F-to-enter,FSWAP same as F with variable interchange,R based on increase in multiple R,RSWAP same as R with variable interchange.(F/prev.) |                                                                                                                                                        |                                                                                                                                          |
|                                                                                    | TYPE=name.                                                                                                                                                                              | ZERO zero intercept, NONZERO, FLOAT treat intercept as a variable to be entered only if significant. (NONZERO/prev.)                                   |                                                                                                                                          |
|                                                                                    | LEVEL=#list.                                                                                                                                                                            | Number list, corresponding to variable order, specifying level for forcing entry. Zeros will not be entered. Lowest level entered first. (highest + 1) |                                                                                                                                          |

# BMDP3R - NONLINEAR REGRESSION

FORCE=#. Level through which variables will be forced into equation. (none/prev.)  
 STEP=#. Maximum steps. (twice # of variables/prev.)

## BMDP3R NONLINEAR REGRESSION

Least squares fit to a non linear function. Five most frequently used non linear functions are built in, others may be specified with FORTRAN statements.

### OUTPUT

- Parameter values and residual sum of squares.
- Asymptotic correlations and std. dev. for est. parameters.
- Predicted and observed value for the dependent variable.
- Residual value.
- Observed values for independent variables.

### SPECIFYING SPECIAL REGRESSION FUNCTIONS IN FORTRAN

Any function can be fitted to the data by specifying the function and its derivatives in the FORTRAN subroutine FUN. The FORTRAN subroutine containing the correct regression function and derivatives must then become part of the program BMDP3R for that problem run. If the derivatives of the desired nonlinear expression are difficult to obtain or specify, the user is advised to use BMDPAR which does not require the derivatives to be specified. The following is an example subroutine FUN for the regression equation

$$f=p(1)*EXP(P(2)*t)+p(3)*EXP(P(4)*t):$$

```
SUBROUTINE FUN(F,DF,P,X,N,KASE,NVAR,NPAR,IPAS,XLOSS)
DIMENSION DF(NPAR),P(NPAR),X(NVAR)
IMPLICIT INTEGER*4 (I-N)
DF(1)=EXP(P(2)*X(3))
DF(2)=P(1)*X(3)*DF(1)
DF(3)=EXP(P(4)*X(3))
F=P(1)*DF(1)+P(3)*DF(3)
RETURN
END
```

### CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                     | PARAMETER                             | COMMENTS                                                                                                        |
|-------------------------------------------------------------------------------|---------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/END<br>/PLOT<br>(Optional) | see MAIN CHART                        |                                                                                                                 |
|                                                                               | VARIAB=list.<br>NORMAL.or<br>DNORMAL. | Variables for plots.(none/prev.)<br>Print normal or detrended normal probability plots of residuals. (no/prev.) |
|                                                                               | SIZE=#1,#2,<br>RESID.                 | Length of X & Y axis.(50,50/prev.)<br>Plot residuals and squared residuals against predicted.(no/prev.)         |
| /REGRES<br>(Required)                                                         | TITLE='text'.<br>NUMBER=#.            | Regression title <160 cols.(blank)<br>Select built-in function.(See manual)                                     |
|                                                                               | DEPEND=name.                          | Dependent variable required. (none/prev.)                                                                       |
|                                                                               | TOL=#.                                | Singularity tolerance.(.001/prev.)                                                                              |
|                                                                               | WEIGHT=name.                          | Variable containing the case                                                                                    |

|            |             |                                                                         |
|------------|-------------|-------------------------------------------------------------------------|
|            | PARAM=#.    | weights.(none/prev.)                                                    |
|            | CONST=#.    | # of parameters in the regression function required.(none/prev.)        |
|            | ITER=#.     | The number of constraints on parameters.(none/prev.)                    |
|            | CONVER=#.   | Max # of iterations.(50/prev.)                                          |
|            | HALV=#.     | Convergence tolerance.(.00001/prev.)                                    |
|            | MEANSQ=#.   | Max # of increment halving.(5/prev.)                                    |
|            | PASS=#.     | # used to compute asymptotic standard deviations.(none/prev.)           |
|            | LOSS.       | # of passes of data at each iteration.(1/prev.)                         |
| /PARAM     | INIT=#list. | Loss function specified in FORTRAN subroutine.(no/prev.)                |
| (Optional) | NAME=names. | Starting values for regression parameters.(0,0,..../prev.)              |
|            | MAX=# list. | Names of regression parameters. Up to 8 characters.(P(subscript)/prev.) |
|            | MIN=# list. | Maximum values for regression parameters.(none/prev.)                   |
|            |             | Lower limits for regression parameters.(none/prev.)                     |

BMDP4R REGRESSION ON PRINCIPLE COMPONENTS

Computes a regression analysis for each dependent variable on a set of principle components computed from the independent variables.

OUTPUT

- Means, std. devs. and correlation or covariance matrix.
- Eigenvalues and eigenvectors (principle component coefficients).
- Regression coefficients for principle components and independent variables.
- Residual sums of squares and F ratios at each step.
- Optional principle component scores, scatter plots of raw data, and/or normal and detrended normal prob. plots of residuals.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                    | PARAMETER                             | COMMENTS                                                                                                      |
|----------------------------------------------|---------------------------------------|---------------------------------------------------------------------------------------------------------------|
| /PROB                                        | } see MAIN CHART                      |                                                                                                               |
| /INPUT                                       |                                       |                                                                                                               |
| /VARIAB                                      |                                       |                                                                                                               |
| /TRANSF                                      |                                       |                                                                                                               |
| /SAVE                                        |                                       |                                                                                                               |
| /END                                         |                                       |                                                                                                               |
| /PRINT<br>(Optional)                         | SCORE.                                | Print standardized principle component scores for each case.(no/prev.)                                        |
| /PLOT<br>(Optional)                          | VARIAB=list.<br>NORMAL.or<br>DNORMAL. | Variables included.(none/prev.)<br>Print normal or detrended normal probability plots of residuals.(no/prev.) |
|                                              | SIZE=#1,#2.                           | Length of X & Y axis.(50,50/prev.)                                                                            |
|                                              | RESID.                                | Plot residuals.(no/prev.)                                                                                     |
| /REGRES<br>(Required,<br>may be<br>repeated) | TITLE='text'.<br>DEPEND=name.         | Regression title <80 cols.(blank)<br>Dependent variable required 1st.<br>(none/prev.)                         |
|                                              | INDEP=list.                           | Independent variable list.(those in USE statement/prev.)                                                      |
|                                              | STAND.                                | Regression on standardized vari-                                                                              |

# BMDP5R - POLYNOMIAL REGRESSION

EIGEN.           ables.(no/prev.)  
                   Enter components by size of  
                   eigenvalues. (no/prev.)  
 CORR.            Enter components by size of abso-  
                   lute value of correlation with  
                   dep.var. (yes/prev.)  
 LIMIT=#1,#2.    Smallest correlation and eigen-  
                   value to cause variable to enter.  
                   (.01,.01/prev.)

## BMDP5R      POLYNOMIAL REGRESSION

Fits a polynomial in one variable to the dependent variable. The form of the regression equation is:

$$y(i)=B(0)+B(1)*x+B(2)*x**2+...+B(r)*x(i)**r+e(i)$$

### OUTPUT

- Correlation matrix for regression coefficient estimates.
- Estimates of coefficients and their std. errors.
- A summary table of goodness-of-fit statistics. Optionally residual and fitted polynomial values, scatter plots of observed and predicted or residual values vs. independent variable, normal and detrended norm prob, plots of residuals, coefficients in the expansion of each polynomial P(x) in powers of x.

### CONTROL LANGUAGE USAGE

| PARAGRAPH                  | PARAMETER             | COMMENTS                                                     |
|----------------------------|-----------------------|--------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB | WEIGHT=v.             | Variable for case weights.(none/prev.)                       |
| /TRANSF<br>/SAVE<br>/END   | see MAIN CHART        |                                                              |
| /PLOT<br>(Optional)        | RREG.                 | Print correlation matrix of reg. coefficients. (no/prev.)    |
|                            | DEGREE=#list.         | Print residuals at stated degrees of regression.(none/prev.) |
|                            | ORTHPOL.              | Print orthogonal poly. coefficients. (no/prev.)              |
| /PLOT<br>(Optional)        | NORMAL.or<br>DNORMAL. | Print normal plots.(no/prev.)                                |
|                            | SIZE=#1,#2.           | Length of X & Y axis.(50,50/prev.)                           |
|                            | DEGREE=#list.         | Print plots at specified degrees of regression.(none/prev.)  |
| /REGRES<br>(Required)      | TITLE='text'.         | Regression title <80 cols.(blank)                            |
|                            | DEPEND=name.          | Dependent variable.(none/prev.)                              |
|                            | INDEP=list.           | Independent variables.(none/prev.)                           |
|                            | DEGREE=#.             | Degree of final regression equation. (3/prev.)               |

## BMDP6R      PARTIAL CORRELATION AND MULTIVARIATE REGRESSION

Computes the partial correlations of a set of variables removing the linear effects of a second set of variables. The program can also be used for regression, especially if multiple dependent variables are present. Computation can be performed in either single or double precision.

### OUTPUT

- Means, std. devs., coeff. of variation, skewness,kurtosis, min.,max.

- Correlations, covariances, partial correlations and partial covariances.
- Regression coef., std. coef., std. errors for coef.
- Covariances and correlations for reg. coefs.
- t-tests and significance levels for reg. coefs.
- Squared multiple correlations and significant levels.
- Residuals from partial correlation analysis.
- Bivariate plots of variables and residuals.

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                         | PARAMETER                                                                                | COMMENTS                                                                                                                               |
|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIABLES<br>/TRANSF<br>/SAVE<br>/END<br>/PRINT<br>(Optional) | } see MAIN CHART<br>TYPE=name.<br>SHAPE=name.<br>CONTENT=list.<br>CASE=#.<br>MATR=names. | One of: DATA, COVA, CORR. (DATA) SQUARE or LOWER for CORR or COVA. (SQUARE/prev.)                                                      |
|                                                                                   |                                                                                          | One or more of DATA, COVA, COVAPART. (DATA)                                                                                            |
|                                                                                   |                                                                                          | Number of cases of data printed. (5/prev.)                                                                                             |
|                                                                                   |                                                                                          | One or more of CORR, COVA, PART, COVAPART, RESI, COEFF, TTEST, STANC std. reg. coefs., RREG correlations of the reg.coefs.(CORR, PART) |
|                                                                                   |                                                                                          | X-axis variables. (none)                                                                                                               |
| /PLOT<br>(Optional)                                                               | YVAR=list.                                                                               | Y-axis variables. (none)                                                                                                               |
|                                                                                   | SIZE=#1, #2.<br>NORMAL.                                                                  | Size of plots. (50,50/prev.)<br>Print normal prob. plots.(no/prev.)                                                                    |
| /REGRES<br>(Required)                                                             | DEPEND=list.                                                                             | Names of dependent variables required. (none)                                                                                          |
|                                                                                   | INDEP=list.                                                                              | Independent variables required. (none)                                                                                                 |
|                                                                                   | TOL=#.                                                                                   | Tolerance for matrix inversion. (.01/prev.)                                                                                            |
|                                                                                   | PREC=name.                                                                               | DOUBLE double precision, SINGLE single precision. (DOUBLE/prev.)                                                                       |
|                                                                                   | ZERO.                                                                                    | Assume zero means.(no/prev.)                                                                                                           |

BMDP9R ALL POSSIBLE SUBSETS REGRESSION

Identifies 'best' subsets of predictor variables. Best is defined in terms of the same R-squared, adjusted R-squared or Mallow's Cp.

OUTPUT

- Means, std. devs., coef. of variation, minimums, maximums, minimum std. scores, maximum std. scores, skewness and kurtosis.
- Correlation and covariance matrix.
- Bivariate plots.
- For best M subsets, selection criteria, regression coef. and F-ratio.
- For best subset, R-squared, F ratio and significance, reg. coefs., std. coefs., std. errors, t-tests, significance and tolerance, etc.

CONTROL LANGUAGE USAGE

| PARAGRAPH | PARAMETER | COMMENTS |
|-----------|-----------|----------|
|-----------|-----------|----------|

## BMDPAR - DERIVATIVE-FREE NON LINEAR REGRESSION

|            |   |                |                                                                                                               |
|------------|---|----------------|---------------------------------------------------------------------------------------------------------------|
| /PROB      | } | SHAPE=name.    | SQUARE, LOWER. (SQUARE)                                                                                       |
| /INPUT     |   | TYPE=name.     | One of: DATA, COVA, CORR. (DATA)                                                                              |
| /VARIAB    |   | WEIGHT=v.      | Variable to weight cases (none/prev.)                                                                         |
| /TRANSF    | } | see MAIN CHART |                                                                                                               |
| /SAVE      |   | CONTENT=list.  | One or both of: DATA, COVA.(DATA)                                                                             |
| /END       |   |                |                                                                                                               |
| /PRINT     |   | CASE=#.        | Number of cases data printed.(5/prev.)                                                                        |
| (Optional) |   | MATR=list.     | One or more of CORR, COVA, RESI, RREG. (CORR/prev.)                                                           |
| /PLOT      |   | XVAR=list.     | X-axis variables. (none)                                                                                      |
| (Optional) |   | YVAR=list.     | Y-axis variables. (none)                                                                                      |
|            |   | NORMAL.        | Normal plots. (no/prev.)                                                                                      |
| /REGRES    |   | DEPEND=name.   | Dependent variables, at least 3. (none)                                                                       |
| (Required) |   | INDEP=list.    | Independent variables, at least 3. (none)                                                                     |
|            |   | METH=name.     | One of: RSQ R-squared, ADJ adj. R-squared, CP Mallows Cp, NONE ordinary regression on all subsets. (CP/prev.) |
|            |   | PEN=#.         | Number to add to Cp as penalty for including a variable.(2/prev.)                                             |
|            |   | NUM=#.         | Number of 'best' subsets to be chosen. (10 max.) (5/prev.)                                                    |
|            |   | TOL=#.         | Tolerance for matrix inversion. (.0001/prev.)                                                                 |
|            |   | ZERO.          | Request zero intercept.(no/prev.)                                                                             |

## BMDPAR DERIVATIVE-FREE NON LINEAR REGRESSION

Estimates the parameters of a non linear function by least squares using a pseudo-Gauss-Newton algorithm (Ralston and Jennrich, 1977). PAR is appropriate for a wide variety of functions that are not linear in the parameters, and for which derivatives are difficult to specify or costly to compute. The non linear function must be specified by FORTRAN statements input into the code of the program prior to running the program. The derivatives are not specified.

### OUTPUT

- Weighted means and std. devs. for each variable.
- Residual sum of squares at each iteration.
- The minimal residual sum of squares and estimates of the parameters after the final iteration.
- Final asymptotic correlation matrix of the parameter estimates.
- Estimated final mean square error.
- The asymptotic std. devs. of the parameter estimates (zeros indicate estimation problems.)
- A table containing, for each case; The residual y-f, the observed y, the predicted value, the standard error of the predicted value and the values of the independent variables and the case WEIGHT variable.

### SPECIFYING THE REGRESSION FUNCTION IN FORTRAN

Any function can be fitted to the data by specifying it in the FORTRAN subroutine FUN. The specific FORTRAN subroutine FUN containing the regression function desired a problem must then become a part of the program BMDPAR for that run. The following is an example subroutine FUN for the regression equation.

$$f=p(1)*EXP(p(2)*t)+p(3)*EXP(p(4)*t):$$

```

SUBROUTINE FUN(F,P,X,N,KASE,NVAR,NPAR,IPASS,XLOSS)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
IMPLICIT INTEGER*4(I-N)
DIMENSION P(NPAR),X(NVAR)
F= P(1)*DEXP(P(2)*X(3))+P(3)*DEXP(P(4)*X(3))
RETURN
END

```

CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                     | PARAMETER                                                                                                                                                                                                                                                                                                                                                                                                     | COMMENTS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |  |  |
|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|
| /PROB<br>/INPUT<br>/VARIAB<br>/TRANSF<br>/SAVE<br>/END<br>/REGR<br>(Required) | } see MAIN CHART<br><br>NUMB=#.<br>DEPEND=v.<br>PARAM=#.<br>TITLE='c'.<br>NUMB=#.<br>WEIGHT=v.<br>CONST=#.<br>PRINT=List.<br>CONV=#.<br>ITER=#.<br>HALV=#.<br>TOL=#.<br>MEANSQ=#.<br>PASS=1.or2.<br>LOSS.<br>INIT=#list.<br>MAX=#list.<br>MIN=#list.<br>NAME=List.<br>DELTA=#list.<br>FIXED=names.<br>or subscripts.<br>CONST=#list.<br>LIMIT=#list.<br>VAR=List.<br>RESID.<br>NORM.<br>DNORM.<br>SIZE=#1,#2. | Number of analyses for which predicted values are saved in BMDP<br>Dependent variable. Required.<br>(none/prev.)<br># of parameters. Required.(none/prev.)<br>Title of analysis, <160 cols.(blank)<br>Code # to select among several regression functions. (none/prev.)<br>Variable containing case weights. (no variable/prev.)<br># of inequality constraints.(none/prev.)<br>Data printed for each case. (all variables/prev.)<br>Convergence criteria.(.00001/prev.)<br>Max. # of iterations.(50/prev.)<br>Max. # of increment halvings. (5/prev.)<br>Check against round-off errors. (10 /prev.)<br>If specified, used to compute asymptotic std. devs.<br># of passes of data for each iteration. (1/prev.)<br>Loss function specified in FORTRAN subroutine. (no/prev.)<br>Starting values for parameters. Required. (none/prev.)<br>Upper limit of parameters.(none/prev.)<br>Lower limits of parameters.(none/prev.)<br>Names for parameters.(P(subscript)/prev.)<br>Step size for computing additional starting values of the function. Rarely needed. (.1* INIT (1) when INIT (1)= 0, .01 otherwise/prev.)<br>Parameters whose estimates are fixed at initial values.(none)<br>Coefficients of inequality constraint, may be repeated.(0,0.../prev.)<br>Upper & lower limits for CONST, may be repeated. (- , /prev.)<br>Variables against which predicted and observed values and residuals are plotted. (none/prev.)<br>Residuals and residuals squared plotted against predicted values. (no/prev.)<br>Print normal probability plot of residuals.(no/prev.)<br>Print detrended normal prob. plot. (no/prev.)<br>Width & height of plot. (50,50/ |  |  |
|                                                                               |                                                                                                                                                                                                                                                                                                                                                                                                               | /PARAM<br>(Required)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |  |
|                                                                               |                                                                                                                                                                                                                                                                                                                                                                                                               | /PLOT<br>(Optional)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |  |

# BMDP1S - MULTIPASS TRANSFORMATIONS

prev.)

## BMDP1S MULTIPASS TRANSFORMATION

Can be used when information from the data is required to compute transformations.

### OUTPUT

For each pass, the transformed data and any statistics computed. The transformed data can be saved on a save file.

### CONTROL LANGUAGE USAGE

| PARAGRAPH                               | PARAMETER      | COMMENTS                                                                     |
|-----------------------------------------|----------------|------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/SAVE     | see MAIN CHART |                                                                              |
|                                         | PASS=#.        | Sequence # of PASS in which to create BMDP save file, required. (none/prev.) |
| /END<br>/PASS<br>(Required)<br>/COMPUTE | TITLE='text'   | Alphanumeric pass title <160 cols. (blank)                                   |
|                                         | MEAN.          | Compute mean and std.dev.(yes/prev.)                                         |
|                                         | HARMONIC.      | Compute harmonic means.(no/prev.)                                            |
|                                         | GEOMETRI.      | Compute geometric means.(no/prev.)                                           |
|                                         | EXTR.          | Compute max. & mins. (no/prev.)                                              |
|                                         | VARIAB=list.   | Variables used in computations. (all variables)                              |
| /PRINT<br>(Optional)                    | MEAN.          | Print means and std. devs.(yes/prev.)                                        |
|                                         | HARMONIC.      | Print harmonic means.(no/prev.)                                              |
|                                         | GEOMETRI.      | Print geometric means.(no/prev.)                                             |
|                                         | EXTR.          | Print max. & mins. (no/prev.)                                                |
|                                         | DATA.          | Print transformed data.(no/prev.)                                            |
|                                         | VARIAB=List.   | Variables involved.(same as COMPUTE)                                         |

## BMDP3S NONPARAMETRIC STATISTICS

Performs one of more of the following nonparametric statistical tests:

- Sign test.
- Wilcoxon signed-ranks test.
- Kendall rank correlation coeff.
- Spearman rank correlation coeff.
- Friedman two-way ANOVA.
- Kendall coef. of concordance.
- Mann-Whitney U test.
- Kruskal-Wallis one-way ANOVA.

### OUTPUT

- Test statistics requested.
- Level of significance or probability of occurrence for each test statistic except Kendall and Spearman rank correlation coef.

### CONTROL LANGUAGE USAGE

| PARAGRAPH                            | PARAMETER | COMMENTS |
|--------------------------------------|-----------|----------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP |           |          |



|            |   |                |                                           |
|------------|---|----------------|-------------------------------------------|
| /TRANSF    | } | see MAIN CHART |                                           |
| /SAVE      |   |                |                                           |
| /END       |   |                |                                           |
| /TEST      |   |                |                                           |
| (Required) |   | TITLE='text'   | Test title <160 cols.(blank)              |
|            |   | VARIAB=list.   | Variables to be tested.(all in USE/prev.) |
|            |   | SIGN.          | Sign test. (no/prev.)                     |
|            |   | WILCOX.        | Wilcoxon signed-ranks.(no/prev.)          |
|            |   | KENDALL.       | Kendall rank correlation coef. (no/prev.) |
|            |   | SPEAR.         | Spearman's rank corr. coef.(no/prev.)     |
|            |   | FRIED.         | Friedman two-way ANOVA.(no/prev.)         |
|            |   | KRUWAL.        | Kruskal-Wallis one-way ANOVA. (no/prev.)  |

BMDP1V ONE-WAY ANALYSIS OF VARIANCE AND COVARIANCE

Performs a one-way ANOVA, or a one-way analysis of covariance if covariates are specified. Group sizes may be unequal. Parallel analysis can be performed using several dependent variables.

OUTPUT

For ANOVA:

- Group means, ANOVA table and/or pairwise t-test for group means and t-tests for contrasts of group means.

For analysis of covariance on each dependent variable:

- Regression coeffs., std. errors and t values.
- Group means, adj. group means, and std. errors.
- ANOVA table with test of slope.
- Pairwise t-test for adjusted group means and t-tests for contrasts.
- Regression coef. for each covariate in each group.

Optional:

- Correlation matrix for the reg. coeffs. and/or adj. group means.
- Scatter plots.
- Group statistics and variance, covariance and/or correlation matrices.

CONTROL LANGUAGE USAGE

| PARAGRAPH  | PARAMETER | COMMENTS                                        |
|------------|-----------|-------------------------------------------------|
| -----      |           |                                                 |
| /PROB      | }         | see MAIN CHART                                  |
| /INPUT     |           |                                                 |
| /VARIAB    |           |                                                 |
| /GROUP     |           |                                                 |
| /TRANSF    |           |                                                 |
| /SAVE      |           |                                                 |
| /END       |           |                                                 |
| /PRINT     |           |                                                 |
| (Optional) | MAX.      | Print max. value of vars. by groups. (no/prev.) |
|            | MIN.      | Print min. value of vars. by groups. (no/prev.) |
|            | MEAN.     | Print mean value in groups.(no/prev.)           |
|            | COVA.     | Print covariance matrix.(no/prev)               |
|            | CORR.     | Print correlation matrix. (no/prev.)            |

# BMDP2V - REPEATED MEASURES ANOVA & ANCOVA

|                                              |               |                                                                                                                         |
|----------------------------------------------|---------------|-------------------------------------------------------------------------------------------------------------------------|
|                                              | TOT.          | Print total covariance matrix.(no/prev.)                                                                                |
|                                              | BET.          | Print between cov. matrix.(no/prev)                                                                                     |
|                                              | WIT.          | Print pooled within groups cov. (matrix. (no/prev.))                                                                    |
|                                              | RREG.         | Print corr. between coefs. of covariates. (no/prev.)                                                                    |
|                                              | MCOR.         | Print corr. matrix for adj. group means. (no/prev.)                                                                     |
|                                              | TITLE='text'. | Subproblem title <160 cols.(blank)                                                                                      |
| /DESIGN<br>(Required,<br>may be<br>repeated) | DEPEND=list   | Dependent variables. (all except GROUPING variable/prev.)                                                               |
|                                              | INDEP=list.   | Independent variables. (none)                                                                                           |
|                                              | GROUP=list.   | Groups in subproblem.(all/prev.)                                                                                        |
|                                              | CONTR=list.   | Coefficients for linear contrasts.                                                                                      |
|                                              | PLOT.         | Plot each dependent variable against the independent variables, residuals, squared residuals, and predicted. (no/prev.) |

## BMDP2V ANALYSIS OF VARIANCE AND COVARIANCE, INCLUDING REPEATED MEASURES

Performs ANOVA or analysis of covariance for a general repeated measures model. For each subject, the trial factors must have a complete factorial structure with no missing observations. The group factors can be crossed or nested; group indices are read as data and determine the nesting relationships without further specification.

Unequal cell sizes are allowed but not empty cells.

### OUTPUT

- ANOVA table.
- Cell means and std. devs. for dep. var. and covariates.
- Adjusted cell means and reg. coef.
- Design information.

### CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                                 | PARAMETER       | COMMENTS                                                               |
|-------------------------------------------------------------------------------------------|-----------------|------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/TRANSF<br>/SAVE<br>/END<br>/DESIGN<br>(Required) | see MAIN CHART  |                                                                        |
|                                                                                           | RESI.           | Save residuals and predicted. (yes/prev.)                              |
|                                                                                           | FORM='c'.       | Alphanumeric design.(none/prev.)                                       |
|                                                                                           | GROUP=list.     | Subscripts of grouping variables. (none/prev.)                         |
|                                                                                           | LEVEL=list.     | Subscripts of dependent variables. (none/prev.)                        |
|                                                                                           | DEPEND=list.    | Subscripts of dependent variables.                                     |
|                                                                                           | COVA=list.      | Subscripts of covariates.(none/prev.)                                  |
|                                                                                           | NAME=list.      | Names of trial factors. Up to 4 characters.(R,S,T,V,W,Y,Z/prev.)       |
|                                                                                           | ORTH.           | Requests orthogonal decomposition of the trial main effects.(NO/prev.) |
|                                                                                           | POINT(#)=#list. | Spacing for levels of trial factor                                     |

SYMM. #, may be repeated. (1,2,3,...)  
 Test for compound symmetry.(no/  
 prev.)  
 INCL=#list. Terms to be included or excluded  
 from the model for the grouping  
 factors. (include all main effects  
 and iterations for grouping fac-  
 tors)  
 RESI=name. Method of computing residuals and  
 predicted values MEAN or ORTH.  
 (MEAN/prev.)  
 PRINT. Residuals and predicted values  
 printed.(no/prev.)  
 TOL=#. Tolerance limit. (.01/prev.)

EXAMPLES OF DESIGN:

| LONG FORM                                                     | SHORT FORM         |
|---------------------------------------------------------------|--------------------|
| GROUP=1. DEPEND=2,3,4. LEVEL=3.                               | FORM='G,3(Y)'      |
| GROUP=1,2. DEPEND=3 TO 8. LEVEL=2,3.                          | FORM='2G,2(3(Y))'  |
| GROUP=1. DEPEND=3,5,7. COVA=2,4,6.<br>LEVEL=3.                | FORM='G,3(X,Y)'    |
| GROUP=1. DEPEND=3,4,6,7,9,10.<br>LEVEL=3,2. COVA=2,2,5,5,8,8. | FORM='G,3(X,2(Y))' |

BMDP3V GENERAL MIXED MODEL ANALYSIS OF VARIANCE

This program uses the maximum likelihood (ML) and re-  
 stricted maximum likelihood (REML) approaches to the fixed  
 and random coefficients model. The program is designed to  
 handle mixed models of quite arbitrary form without  
 requiring the balance demanded by P2V. On the other hand,  
 the maximum likelihood approach is not as well studied or  
 as widely available as the more classical approach of P2V.  
 Therefore, it is recommended that this program be used in  
 conjunction with more classical approaches and for mixed  
 model problems not satisfying the balance requirements of  
 P2V or BMD0BV.

OUTPUT

- Reports the estimates and std. devs. of the fixed effects parameters, variance components, estimates of the cell means, and residuals.
- The log-likelihoods for the model and for sub models (that are specified) are computed.

CONTROL LANGUAGE USAGE

| PARAGRAPH                  | PARAMETER                     | COMMENTS                                                                             |
|----------------------------|-------------------------------|--------------------------------------------------------------------------------------|
| /PROB<br>/INPUT<br>/VARIAB | } see MAIN CHART<br>WEIGHT=V. | Variable containing weights.(none/<br>prev.)                                         |
| /GROUP                     |                               |                                                                                      |
| /TRANSF                    |                               |                                                                                      |
| /SAVE                      |                               |                                                                                      |
| /END                       |                               |                                                                                      |
| /DESIGN<br>(Required)      | DEPEND=v.                     | Dependent variable, required.(none/<br>prev.)                                        |
|                            | COVA=List.                    | Covariates.(none/prev.)                                                              |
|                            | FIXED=List.                   | Variables defining fixed effects<br>components, may be repeated.(none/<br>prev.)     |
|                            | FNAME=List.                   | Names of fixed components.(FIX<br>(subscript)/prev.)                                 |
|                            | RAND=List.                    | Variables that define a random<br>effects component, may be repeated<br>(none/prev.) |

# BMDP1L - LIFE TABLE ANALYSIS

|                       |              |                                                                |
|-----------------------|--------------|----------------------------------------------------------------|
|                       | METH=name.   | One only of ML or REML.                                        |
|                       | PARAM=#list. | Initial estimates for parameters.<br>(all 0.0 except =1/prev.) |
|                       | MAXIT=#.     | Max. # of iterations.(30/prev.)                                |
| /HYPOTH<br>(Optional) | FIX=#list.   | Fixed components set to zero.<br>(none)                        |
|                       | RAND=#list.  | Random components set to zero.<br>(none)                       |
| /GROUP                |              | Required for fixed effects. See<br>MAIN CHART.                 |

## BMDP1L LIFE TABLES AND SURVIVAL FUNCTIONS

Estimates the survival (time-to-response) distribution of patients who have been observed over varying periods of time. The data can arise from experiments in which all subjects enter the study either concurrently or at different times. In the latter case, all patients are observed for unequal amounts of time. A major advantage of the techniques used in this program is that all patient follow-up information (although unequal in time) is used in estimating the survival probabilities.

### OUTPUT

- Estimates of the survival function using the actuarial-type life table(Cutler-Ederer).
- Estimates of the survival function based on the product limit (Kaplan-Meier) estimate.
- The Mantel-Cox and Breslow statistics to test the equality of survival distributions for different patient groups (these statistics are the analogues of non parametric rank statistics that are appropriate to censored data.)

### CONTROL LANGUAGE USAGE

| PARAGRAPH                                                                                                     | PARAMETER                     | COMMENTS                                                                               |                                                                                                                                               |
|---------------------------------------------------------------------------------------------------------------|-------------------------------|----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| -----                                                                                                         | -----                         | -----                                                                                  |                                                                                                                                               |
| /PROB<br>/INPUT<br>/VARIAB<br>/GROUP<br>/TRANSF<br>/SAVE<br>/END<br>/EST<br>(Optional)<br>may be<br>repeated) | } see MAIN CHART<br><br>..... | Required if GROUPing var. is<br>specified and it has more than<br>10 different values. |                                                                                                                                               |
|                                                                                                               |                               | METH=name.<br>PRINT.                                                                   | One only of LIFE,PROD.(LIFE/prev.)<br>Print survival distribution estimate.                                                                   |
|                                                                                                               |                               | one PERIOD=#.<br>only WIDTH=#.<br>CUTP=#list.<br>VAR=list.                             | Provide the intervals for the life<br>table when METH=LIFE. (PERIOD=10<br>prev.)<br>Variables printed with product-<br>limit estimate. (none) |
|                                                                                                               |                               | PLOT=list.                                                                             | One or more of SURV,LOG,CUM,HAZ,<br>DEN. (SURV/prev.)                                                                                         |
|                                                                                                               |                               | SIZE=#1,#2.                                                                            | Width and height of plots.(100,<br>50/prev.)                                                                                                  |
|                                                                                                               |                               | GROUP=v.                                                                               | Variable used to classify cases<br>into groups.(none)                                                                                         |
| /FORM                                                                                                         | STAT=list.                    | One or both of MANTEL, BRESLOW.<br>(none/prev.)                                        |                                                                                                                                               |
|                                                                                                               | UNIT=name.                    | One only of DAY, WEEK, MONTH,<br>YEAR. (MONTH)                                         |                                                                                                                                               |
|                                                                                                               | TIME=v.                       | Variable that contains survival<br>times in UNITS. (none)                              |                                                                                                                                               |
|                                                                                                               | STAT=v.                       | Variable containing status of<br>patient. (none)                                       |                                                                                                                                               |
|                                                                                                               | RESP=#list.                   | Values of STATus variable that<br>are responses.(first code of<br>STATus)              |                                                                                                                                               |
|                                                                                                               | LOSS=#list.                   | Values of STATus variable that                                                         |                                                                                                                                               |

ENTRY=List. are losses.(none)  
 Month, day and year of patient  
 entry. (none)  
 TERM=List. Month, day and year of patient  
 termination. (none)

The following are used when a life table is used as input.

NEN=v. Variable containing # of patients  
 entering interval.(sum of NDEAD +  
 NLOST + NWITH)  
 NDEAD=v. Variable containing # of patients  
 dying in interval.(none)  
 NLOST=v. Variable containing # lost in  
 interval. (none)  
 NWITH=v. Variable containing # withdrawn  
 during interval.(none)  
 INTER=v. Variable containing lower limits  
 for intervals. (0,1,2,...)

BMDP EXAMPLES=====

EXAMPLE 1:

```

/PROBLEM TITLE IS 'VARIABLE DEFINITIONS AND SAVE FILE CREATION
EXAMPLE USING BMDP1D'.

/INPUT VARIABLES ARE 16.
CASES ARE 50.
FORMAT IS '(F4.0, 2F4.1, 2F4.0, F4.1, 3F4.0, 5F4.0, A4)'.

/VARIABLE LABEL IS ID.
NAMES ARE HREATIN, CALCIUM, PHOSPHRS, CLUCOSE, BUN,
URICACID, CHOLSTR1, PROTEIN, ALBUMIN,
BILIRUBNM, ALKAPHOS, LACTDEHY, SGOT, BRTHPILL,
ILLNESS, ID.
MIN IS (7) 100.0. MAX IS (7) 500.0.

/CATEGORY CUTPOINTS (1) ARE 1.5, 2.5.

/PRINT MIN. MAX.

/SAVE NEW. UNIT IS 9. CONTENT = DATA. CODE = WERNER.
LABEL IS 'WERNER DATA AS OF 3-13-74'.

/END.
.
. 50 data records
.
/FINISH.

```

EXAMPLE II:

```

/PROBLEM TITLE IS 'EXAMPLE T-TEST ON SELECTED GROUPS USING
BMDP3D'.

/INPUT VARIAB ARE 2. CASES ARE 13. FORM IS '(2F2.0)'.

/VARIAB NAMES ARE CATEGORY, VALUE. GROUP = CATEGORY.

/GROUP CODES = 1 TO 4. NAMES ARE A, B, C, D.

/TEST GROUPS ARE A, B, D.

/END.
.
. 13 data records
.
/FINISH.

```

EXAMPLE III:

## BMDP - EXAMPLES

```
/PROBLEM TITLE IS 'DISCRIMINANT ANALYSIS ON THREE GROUPS USING
BMDP7M'.
/INPUT VARIAB = 5. CASES = 150. FORM = '(4F3.1, F3.0)'.
 GROUPS ARE 3.
/VARIABLE NAMES ARE 'SEPAL L', 'SEPAL W', 'PETAL L', 'PETAL W',
 IRIS.
 GROUPING IS IRIS.
/GROUP CODES ARE 1 TO 3.
 NAMES ARE SETOSA, VERSICOL, VIGINIC.
/PRINT CLASSIFICATION IS 1 TO 4.
/END.
 . 150 records
 .
/FINISH.
```

---

### EXAMPLE IV:

```
/PROBLEM TITLE 'FREQUENCY COUNTS USING BMDP2D ON CARD DATA IN
SEPARATE FILE'.
/INPUT UNIT IS 7.
 FORMAT IS '(F4.0, 2F4.1)'.
/VARIABLES NAMES ARE HREATING, CALCIUM, PHOSPHRS.
/END.
/FINISH.
```

---

### EXAMPLE V:

```
/PROBLEM TITLE 'STEPWISE REGRESSION BMDP2R USING COVARIANCE
MATRIX ON A SAVE FILE.'.
/INPUT UNIT IS 9.
 CONTENT IS COVA.
 CODE IS WERNER.
 LABEL IS 'WERNER COVARIANCE MATRIX FROM BMDP8D'.
/PRINT COVARIANCE.
/REGRES TITLE IS 'PREDICT BUN'.
 DEPEND IS BUN.
/END.
/FINISH.
```

---

# INDEX

- ACCEPT (SIR), 27
  - ACCEPT REC IF (SIR), 12
  - across record functions (SIR), 5
  - ADD (SIR), 27
  - ADD REC (SIR), 18
  - AFTER REPORT (SIR), 37
  - AFTER RETRIEVAL (SIR), 21
  - AGGREGATE (SPSS), 49
  - algebraic expressions (SIR), 12
  - altering a schema (SIR), 17
  - AMOD (SIR), 4
  - ANOVA (SPSS), 50
  - ANOVA-mixed model (BMDP), 103
    - one way (BMDP), 101
    - repeated measures (BMDP), 102
    - SPSS, 50
  - arithmetic expressions (SIR), 12
  - arithmetic operators (SIR), 3
  - ASSIGN MISSING (SIR), 21
  - assigning password (SIR), 9
  - assigning security (SIR), 9
  - AT END (SIR), 37
- 
- batch retrieval commands (SIR), 20
  - batch retrieval example (SIR), 42
  - BEFORE REPORT (SIR), 38
  - bivariate plotting (BMDP), 78
  - BLANK LINES (SIR), 21
  - block clustering (BMDP), 86
  - blocks in retrievals (SIR), 20
  - BMDP SAVE FILE (SIR), 34
  - BMDP-computation (BMDP), 74
    - control language (BMDP), 68
    - examples (BMDP), 105
    - introduction (BMDP), 68
    - main chart (BMDP), 70
    - reserved words (BMDP), 69
    - statistical programs (BMDP), 75
    - transformations (BMDP), 74
  - BMDP1D-simple data desc. (BMDP), 75
  - BMDP1F-two-way contingency (BMDP), 81
  - BMDP1L-life tables (BMDP), 104
  - BMDP1M-cluster analysis (BMDP), 85
  - BMDP1R-linear regression (BMDP), 91
  - BMDP1S-multipass trans. (BMDP), 100
  - BMDP1V-one way ANOVA (BMDP), 101
  - BMDP2D-frequency counts (BMDP), 76
  - BMDP2F-two-way with indep. (BMDP), 83
  - BMDP2M-cluster analysis (BMDP), 86
  - BMDP2R-stepwise m.l. reg. (BMDP), 92
  - BMDP2V-repeated meas. ANOVA (BMDP), 102
  - BMDP3D-t-tests (BMDP), 76
  - BMDP3F-multiway cont. tables (BMDP), 84
  - BMDP3M-block clustering (BMDP), 86
  - BMDP3R-nonlinear regression (BMDP), 94
  - BMDP3S-nonparametric stats (BMDP), 100
  - BMDP3V-mixed model ANOVA (BMDP), 103
  - BMDP4D-alphanumeric counts (BMDP), 77
  - BMDP4M-factor analysis (BMDP), 87
  - BMDP4R-principle comp. reg. (BMDP), 95
  - BMDP5D-univariate plots (SPSS), 78
  - BMDP5R-polynomial reg. (BMDP), 96
  - BMDP6D-bivariate plots (BMDP), 78
  - BMDP6M-canonical corr. (BMDP), 88
  - BMDP6R-multivariate reg. (BMDP), 96
  - BMDP7D-histograms & ANOVA (BMDP), 79
  - BMDP7M-discriminant analysis (BMDP), 89
  - BMDP8D-correlation (BMDP), 80
  - BMDP9D-multidimensional data (BMDP), 81
  - BMDP9R-all subsets reg. (BMDP), 97
  - BMDPAM-missing data est. (BMDP), 90
- 
- BMDPAR-non linear reg. (BMDP), 9
  - branching (SIR), 33
  - BREAKDOWN (SPSS), 50
- 
- CALL (SIR), 32
  - CANCORR (SPSS), 51
  - canonical correlation (BMDP), 8
  - case accessing (SIR), 22
  - case definition (SIR), 9
  - case deletion (SIR), 32
  - CASE ID (SIR), 9
  - CASE IS, (SIR), 32
  - case selection (SIR), 32
  - case structure listing (SIR), 3
  - CDATE (SIR), 3
  - changing a schema (SIR), 17.
  - character set (SIR), 1
  - CHECK (SIR), 178
  - CIR (SIR), 10
  - CLEAR (SIR), 27
  - CNT (SIR), 4
  - CNTR (SIR), 5
  - command format (SIR), 2
  - COMMENT (SIR), 6
  - comments in SIR (SIR), 6
  - common information record (SIR), 10
  - COMMON LIST (SIR), 10
  - computation (BMDP), 74
  - COMPUTE (SIR), 12, 22
  - CONDESCRIPTIVE (SPSS), 52
  - CONDESCRIPTIVE (SIR), 34
  - contingency table analysis (BMDP), 82
  - continuation lines (SIR), 2
  - control language (BMDP), 68
  - control words (SPSS), 48
  - COPY (SIR), 28
  - COUNT (SIR), 6
  - CROSSTABS (SPSS), 52
  - CTIME (SIR), 3
  - CWRITE (SIR), 28
- 
- DATA LIST (SIR), 12
  - data types (SIR), 2
  - DATE (SIR), 37
  - date functions (SIR), 3
  - DATE VAR LIST (SIR), 12
  - DATET (SIR), 3
  - DELETE (SIR), 28
  - DELETE CASE (SIR), 19, 32
  - DELETE CASE IF (SIR), 19
  - DELETE REC (SIR), 19
  - DELETE REC IF (SIR), 19
  - DELETE RECORD (SIR), 32
  - delimiters (SIR), 1
  - designing retrievals (SIR), 20
  - DETAIL (SIR), 38
  - detailed schema listing (SIR), 1
  - DISCRIMINANT (SPSS), 53
  - DISK RESTORE (SIR), 9
  - DO REPEAT (SIR), 8
  - DOCUMENT (SIR), 6
  - dumping SIR records (SIR), 41
  - dumping to cards (SIR), 36
  - dumping to tape (SIR), 39
- 
- EDIT (SIR), 8
  - editor return (SIR), 30
  - editor usage (SIR), 26
  - EDITOR \$ (SIR), 31
    - ACCEPT (SIR), 27
    - ADD (SIR), 27

INDEX (cont.)

- EDITOR (SIR) cont.  
 -CALL (SIR), 32  
 -CLEAR (SIR), 37  
 -COPY (SIR), 28  
 -CWRITE (SIR), 28  
 -DELETE (SIR), 28  
 -END (SIR), 28  
 -EP (SIR), 31  
 -FINISH (SIR), 28  
 -GET (SIR), 28  
 -HELP (SIR), 28  
 -LIST (SIR), 28  
 -OUTPUT (SIR), 29  
 -PAGE (SIR), 29  
 -PREAD (SIR), 29  
 -PURGE (SIR), 29  
 -PWRITE (SIR), 29  
 -QUIT (SIR), 28  
 -READ (SIR), 29  
 -REPLACE (SIR), 29  
 -RESEQUENCE (SIR), 30  
 -REWIND (SIR), 30  
 -RUN (SIR), 30  
 -SAVE (SIR), 30  
 -TAB (SIR), 31  
 -WRITE (SIR), 31  
 -X (SIR), 31  
 END (SIR), 22,28,38  
 END LEVEL (SIR), 38  
 END REPEAT (SIR), 8  
 END REPORT (SIR), 38  
 EP (SIR), 31  
 ERROR LIMIT (SIR), 8  
 EVICT REC (SIR), 20  
 examples-BMDP (BMDP),105  
 -REPORT (SIR), 43  
 -SIR batch retrieval (SIR),42  
 -SIR schema defn.(SIR),41  
 -SPSS (SPSS), 66  
 execution options (SIR), 45  
 execution parameters (SIR), 27  
 EXISTS (SIR), 6  
 EXIT (SIR), 32,38  
 EXIT BLOCK IF (SIR), 22  
 EXIT IF (SIR), 38
- 
- FACTOR (SPSS), 54  
 factor analysis (BMDP), 87  
 FILE NAME (SIR), 8  
 file statistics (SIR), 16  
 FINISH (SIR), 8,28  
 fixed decimals (SIR), 14  
 fixed format (SIR), 2  
 FOOTING (SIR), 22  
 FOR EACH CASE (SIR), 22  
 FOR EACH REC (SIR), 22  
 formats (SIR), 13  
 formatting (SIR), 25  
 formatting commands (SIR),22  
 FREQUENCIES (SPSS), 44  
 FREQUENCIES (SIR), 34  
 FST (SIR), 4  
 FSTR (SIR), 5  
 functions (SIR), 3
- 
- G3SLS (SPSS), 55  
 general rules (SIR), 1  
 GET (SIR), 28  
 GET FILE (SIR), 8  
 GUTTMAN SCALE (SPSS) ,55
- 
- HEADING (SIR), 21,23  
 HELP (SIR), 28  
 histograms (SIR), 34  
 histograms (SPSS), 54  
 histograms (BMDP), 78
- 
- IF (SIR), 12, 23, 32  
 index key size (SIR), 10  
 initialization (SIR), 2  
 initializing variables (SIR), 23  
 INPUT FORMAT (SIR), 13  
 INPUT MEDIUM (SIR), 7  
 input record size (SIR), 10  
 integer constants (SIR), 2  
 INTEGER LIST (SIR), 13  
 integer variables (SIR), 13  
 interactive retrieval example (SIR), 44  
 interactive usage (SIR), 26  
 introduction (BMDP), 68
- 
- JFACTOR (SPSS), 56  
 JULC (SIR), 4  
 JULN (SIR), 4  
 JUMP (SIR), 33  
 JUMP TO LEVEL, 38
- 
- key size (SIR), 10
- 
- labels listing (SIR), 16  
 labels for variables (SIR), 25  
 LEN (SIR), 4  
 LEVEL (SIR), 38  
 LIFE TABLES (BMDP), 104  
 linear regression (SPSS), 61  
 linear regression (BMDP), 91  
 LIST (SIR), 28  
 LIST SCHEMA (SIR), 16  
 LIST STATS (SIR), 16  
 listing SIR files (SIR), 41  
 locked schema (SIR), 17  
 logical computation (SIR),12  
 logical operators (SIR), 3  
 LSTR (SIR), 5
- 
- main chart (BMDP), 70  
 MANOVA (SPSS), 57  
 MAX (SIR), 4  
 MAX INPUT COLS (SIR), 10  
 MAX KEY SIZE (SIR), 10  
 MAX REC COUNT (SIR), 10, 13  
 MAX REC TYPES (SIR), 10  
 MAXR (SIR), 5  
 MEAN (SIR), 4  
 MEANR (SIR), 5  
 merging SIR files (SIR),40  
 MIN (SIR), 5  
 MINR (SIR), 5  
 MISSING VALUES (SIR), 13  
 MISSING VALUES-retrievals (SIR), 21  
 mixed model ANOVA (BMDP), 103  
 MODIFY CASE (SIR), 19  
 MODIFY CASE IF (SIR), 19  
 MODIFY REC (SIR), 19  
 MODIFY REC IF (SIR), 19  
 MODIFY SCHEMA (SIR), 19  
 modifying a schema (SIR),17



INDEX (cont.)

- MOVE (SIR), 28
  - MOVE VAR LIST (SIR), 23
  - MULT RESPONSE (SPSS), 58
  - mult.lin.regression (SPSS), 61
  - mult.lin.regression (BMDP), 91
  - multi-pass trans. (BMDP), 100
  - multi-way tables (BMDP), 84
  - multiple summary records (SIR), 22
- 
- N OF CASES (SIR), 11
  - n-way tables (SPSS), 52
  - non linear regression (SPSS), 58
  - non parametric tests (SPSS), 59
  - NONLINEAR (SPSS), 58
  - NONPAR CORR (SPSS), 59
  - NPAR TESTS (SPSS), 59
  - NREAD (SIR), 33
  - NUMBERED (SIR), 9
  - NUMBR (SIR), 4
- 
- ON ERROR-report (SIR), 39
  - ON PAGE -report (SIR), 39
  - one way ANOVA & COVA (SPSS), 50
  - one way ANOVA & COVA (BMDP), 101
  - ONEWAY ANOVA (SPSS), 60
  - order of retrieval commands (SIR), 21
  - order of schema commands (SIR), 11
  - OUTPUT (SIR), 29
  - output specifications (SIR), 25
- 
- PAGE (SIR), 29, 37
  - PAGE EJECT (SIR), 23
  - PAGESIZE (SIR), 9
  - paragraphs (BMDP), 69
  - PARTIAL CORR (SPSS), 61
  - PASSWORD (SIR), 9
  - PEARSON CORR (SPSS), 61
  - PERFORM PROCS (SIR), 22
  - PLOT (SIR), 35
  - plotting-bivariate (BMDP), 78
  - plotting-univariate (BMDP), 78
  - polynomial regression (BMDP), 96
  - PREAD (SIR), 29
  - PRESET (SIR), 23
  - PRINT BACK (SIR), 7
  - procedure calls (SIR), 32
  - procedures (SIR), 34
  - process rec (SIR), 23
  - programming retrievals (SIR), 20
  - prompting (SIR), 33
  - PURGE (SIR), 29
  - PURGE SIR FILE (SIR), 40
  - PWRITE (SIR), 29
- 
- query commands (SIR), 26
  - QUIT (SIR), 28
- 
- RAND (SIR), 6
  - random numbers (SIR), 6
  - READ (SIR), 29
  - READ INPUT DATA (SIR), 18
  - READ SECURITY (SIR), 9
  - real numbers (SIR), 2
  - REC SECURITY (SIR), 14
  - RECODE (SIR), 24
  - record accessing (SIR), 22
  - record definition (SIR), 11
  - record deletion (SIR), 32
  - RECORD IS (SIR), 33
  - record ordering (SIR), 14
  - record processing (SIR), 23
  - RECORD SCHEMA (SIR), 13
  - record selection (SIR), 24
  - record size (SIR), 10
  - RECS PER CASE (SIR), 11
  - RECTYPE COLS (SIR), 11
  - REGRESSION (SPSS), 61
  - regular schema listing (SIR), 16
  - REJECT CASE IF (SIR), 24
  - REJECT REC IF (SIR), 12, 24
  - relational operators (SIR), 3
  - RELIABILITY (SPSS), 62
  - repeated measures ANOVA (BMDP), 102
  - repeating commands (SIR), 8
  - repeating summary records (SIR), 22
  - REPLACE (SIR), 29
  - REPLACE REC (SIR), 18
  - report generator (SIR), 36
  - REPORT-AFTER REPORT (SIR), 37
    - AT END (SIR), 37
    - BEFORE REPORT (SIR), 38
    - DETAIL (SIR), 38
    - END (SIR), 38
    - END LEVEL (SIR), 38
    - END REPORT (SIR), 38
    - EXAMPLE (SIR), 43
    - EXIT (SIR), 38
    - EXIT IF (SIR), 38
    - JUMP TO LEVEL (SIR), 38
    - LEVEL (SIR), 38
    - ON ERROR (SIR), 39
    - ON PAGE (SIR), 39
  - required commands (SIR), 6
  - RESEQUENCE (SIR), 30
  - RETRIEVAL (SIR), 24
  - retrieval blocks (SIR), 20
  - retrieving variables (SIR), 23
  - RETURN (SIR), 30
  - REWIND (SIR), 30
  - reworking a schema (SIR), 17
  - RUN (SIR), 30
  - RUN NAME (SIR), 9
- 
- SAVE (SIR), 30
  - save file -BMDP (SIR), 34
    - SPSS (SIR), 36
  - scaled integers (SIR), 2
  - SCALED VAR LIST (SIR), 14
  - SCATTERGRAM (SPSS), 63
  - schema listing (SIR), 16
  - schema case definition (SIR), 9
  - schema modification (SIR), 16
  - schema record definition (SIR), 1
  - SECURITY (SIR), 9
  - security-records (SIR), 14
    - variables (SIR), 15
  - SELECT CASE IF (SIR), 24
  - SELECT FIRST (SIR), 25
  - SELECT REC IF (SIR), 24
  - sequence check (SIR), 14
  - SEQUENCE COLS (SIR), 11
  - setting page size (SIR), 9
  - setting tabs (SIR), 31
  - setting variable value (SIR), 23
  - SIR FILE DUMP (SIR), 41
  - SIR FILE LIST (SIR), 41
  - SIR MERGE (SIR), 40
  - SIR SAVE FILE (SIR), 35
  - SIR SUBSET (SIR), 40
  - SIR-arithmetic operators (SIR), 3
    - batch retrieval (SIR), 20
    - character set (SIR), 1

INDEX (cont.)

SIR (cont.)

-data entry/update (SIR),18  
 -data types (SIR), 2  
 -editor (SIR), 26  
 -examples (SIR), 4  
 -execution options (SIR),45  
 -functions (SIR), 3  
 -general (SIR), 1  
 -interactive (SIR), 26  
 -REPORT (SIR), 36  
 -schema definition (SIR),9,11  
 -schema modification (SIR),16  
 -schema unlocking (SIR), 17  
 -SECURITY (SIR), 9  
 -special commands (SIR), 32  
 -supplied procedures (SIR), 34  
 -system initialization (SIR), 8  
 -system wide commands (SIR),6  
 -utilities (SIR), 39  
 -variable initialization (SIR),2  
 -variable names (SIR), 2  
 SORT RECORDS (SIR), 14  
 SPACE (SIR), 7  
 special functions (SIR),6  
 special retrieval commands (SIR),  
 31  
 SPECTRAL (SPSS), 63  
 SPSS SAVE FILE (SIR), 36  
 SPSS-AGGREGATE (SPSS), 49  
 -ANOVA (SPSS), 50  
 -BREAKDOWN (SPSS), 50  
 -CANCORR (SPSS), 51  
 -CONDESCRIPTIVE (SPSS),52  
 -control words (SPSS), 48  
 -CROSSTABS (SPSS), 52  
 -DISCRIMINANT (SPSS), 53  
 -examples (SPSS), 65  
 -FACTOR (SPSS), 54  
 -FREQUENCIES (SPSS),54  
 -G3SLS (SPSS), 55  
 -general rules (SPSS), 47  
 -GUTTMAN SCALE (SPSS), 55  
 -JFACTOR (SPSS), 56  
 -MANOVA (SPSS), 57  
 -MULT RESPONSE (SPSS), 58  
 -mult.lin.regression (SPSS),61  
 -multivariate analysis (SPSS),57  
 -non parametric tests (SPSS), 59  
 -NONLINEAR (SPSS), 58  
 -NONPAR CORR (SPSS), 59  
 -ONEWAY (SPSS), 60  
 -PARTIAL CORR (SPSS), 61  
 -PEARSON CORR (SPSS), 61  
 -rank order corr (SPSS), 59  
 -REGRESSION (SPSS), 61  
 -RELIABILITY (SPSS), 62  
 -reserved words (SPSS), 47  
 -SCATTERGRAM (SPSS), 63  
 -SPECTRAL (SPSS), 63  
 -summary statistics (SPSS),52  
 -SUMMARY TABLES (SPSS), 65  
 -SURVIVAL (SPSS), 65  
 SRCH (SIR), 5  
 statistics on the data set (SIR),16  
 STDEV (SIR), 5  
 STDEVW (SIR), 5  
 stepwise mult.lin.reg. (BMDP),92  
 stepwise mult. lin.reg.(SPSS),61  
 storing SIR procedures (SIR),30  
 storing SIR to tape (SIR), 39  
 string functions (SIR), 4  
 STRING LENGTH (SIR), 7  
 string variables (SIR), 7  
 strings (SIR), 3

structure for retrievals (SIR),20  
 structure listing (SIR), 16  
 SUM (SIR), 5  
 SUMMARY (SIR), 25  
 summary statistics (SIR), 34  
 summary statistics (SPSS), 52  
 SUMMARY TABLES (SPSS), 65  
 SUMR (SIR), 5  
 supplied procedures (SIR), 34  
 SURVIVAL (SPSS), 65  
 syntax checking (SIR), 8  
 system initialization (SIR), 8  
 system wide commands (SIR), 6

TABS (SIR), 31  
 TAPE FILE (SIR), 8  
 TAPE STORE (SIR), 39  
 TASK NAME (SIR), 7  
 TEMP VAR LIST (SIR), 11  
 temporary variables (SIR), 11  
 terminal usage (SIR), 26  
 TETRACHORIC (SPSS), 65  
 TIME (SIR), 4,37  
 time functions (SIR), 3  
 time numbers(SIR), 3  
 TIME VAR LIST (SIR), 14  
 transformations (BMDP), 74  
 two way cont. table analysis  
 (BMDP), 82  
 two way frequency tables (SPSS),  
 52

unlocking a schema (SIR), 17  
 UPDATE REC (SIR), 19  
 updating commands (SIR), 18  
 utilities (SIR), 39

VALID VALUES (SIR), 14  
 VALLAB (SIR), 6  
 value label retrieval (SIR),6  
 VALUE LABELS (SIR), 15, 25  
 value recoding (SIR), 24  
 VAR LABELS (SIR), 15, 25  
 VAR RANGES (SIR), 15, 25  
 VAR SECURITY (SIR), 15  
 variable definition (SIR),12,15  
 variable initialization (SIR),  
 2, 23  
 variable label retrieval (SIR),  
 6  
 VARIABLE LIST (SIR), 15  
 variable names (SIR), 2  
 VARLAB (SIR), 6

WARNING LIMIT (SIR), 8  
 within record functions (SIR), 4  
 WRITE (SIR), 25, 31  
 WRITE IF (SIR), 25  
 WRITE RECORDS (SIR), 36  
 WRITE SCHEMA (SIR), 14  
 WRITE SECURITY (SIR), 9

X command (SIR), 17, 31

