AD-A035 754 LETTERMAN ARMY INST OF RESEARCH SAN FRANCISCO CALIF F/G 12/1

GENERALIZED RESEARCH ANALYSIS STATISTICAL SYSTEM. SECOND EDITIO--ETC(U)

NOV 76 H LAZARUS, T MCCAA, R TEPLICK

UNCLASSIFIED LAIR-33 NL

1 OF 2

ADA035754

ADA035754

COPY AVAILABLE TO DDC DOES NOT
PERMIT FULLY LEGIBLE PRODUCTION

14

INSTITUTE REPORT 33

GENERALIZED RESEARCH ANALYSIS STATISTICAL SYSTEM

SECOND EDITION (DECEMBER 1974)

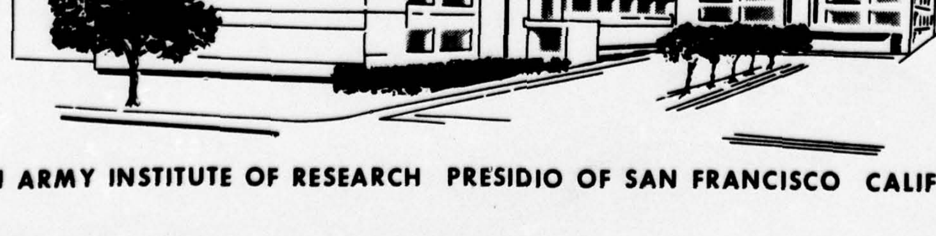
HENRY LAZARUS, SP5
TOM McCAA, SP5
RICHARD TEPLICK, MAJ
MARGARET WRENSCH, M.S.

DEPARTMENT OF INFORMATION SCIENCES
NOVEMBER 1976

D D C
RECEIVED
FEB 18 1977
A

...vailable to DDC does not
...fully legible reproduction

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited



LETTERMAN ARMY INSTITUTE OF RESEARCH PRESIDIO OF SAN FRANCISCO CALIFORNIA 94129

REPRODUCTION OF THIS DOCUMENT IN WHOLE OR IN PART IS PROHIBITED EXCEPT WITH THE PERMISSION OF LETTERMAN ARMY INSTITUTE OF RESEARCH, PRESIDIO OF SAN FRANCISCO, CALIFORNIA 94129. HOWEVER, DDC IS AUTHORIZED TO REPRODUCE THE DOCUMENT FOR UNITED STATES GOVERNMENT PURPOSES.

DESTROY THIS REPORT WHEN NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

THE OPINIONS OR ASSERTIONS CONTAINED HEREIN ARE THE PRIVATE VIEWS OF THE AUTHORS AND ARE NOT TO BE CONSTRUED AS OFFICIAL OR AS REFLECTING THE VIEWS OF THE DEPARTMENT OF THE ARMY OR THE DEPARTMENT OF DEFENSE.

CITATION OF TRADE NAMES IN THIS REPORT DOES NOT CONSTITUTE AN OFFICIAL ENDORSEMENT OR APPROVAL OF THE USE OF SUCH ITEMS.

Copy available to DDC does not permit fully legible reproduction

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) LAIR-33	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) GENERALIZED RESEARCH ANALYSIS STATISTICAL SYSTEM. SECOND EDITION (DECEMBER 1974).		5. TYPE OF REPORT & PERIOD COVERED (9) FINAL rept. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) [redacted] Henry Lazarus, [redacted] Tom McCaa, Richard Teplick [redacted] Margaret Wrensch, M.S.		8. CONTRACT OR GRANT NUMBER(s) (16)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Information Sciences (SGRD-ULI) Letterman Army Institute of Research Presidio of San Francisco, CA 94129		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project #3A762760A822 Work Unit #082
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Medical Research and Development Command, Washington, DC 20314		12. REPORT DATE (11) Nov [redacted]76 13. NUMBER OF PAGES 148
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 153p.		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Computers
Statistics

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report describes an automated Generalized Research Analysis Statistical System (GRASS) which is easy to use and which will satisfy many of the statistical computational requirements of laboratory investigators. Instructions are given for utilization of this system in a manner sufficiently detailed to be usable by personnel unfamiliar with computer operations.

TABLE OF CONTENTS 10/07/76

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

1.1 INTRODUCTION

NEARLY EVERY INVESTIGATOR HAS A NEED, AT ONE TIME OR ANOTHER, FOR STATISTICAL COMPUTATIONS. HE MAY USE A SMALL DESK CALCULATOR TO COMPUTE MEANS AND STANDARD DEVIATIONS FOR EACH OF HIS VARIABLES. HOWEVER, AS THE COMPLEXITY OF THE ANALYSES INCREASES OR AS THE NUMBER OF VARIABLES AND/OR THE NUMBER OF OBSERVATIONS INCREASES, MANY HOURS HAVE TO BE SPENT BEHIND THE CALCULATOR. THIS INCREASES THE CHANCE FOR HUMAN ERRORS IN THE CALCULATIONS. IN ADDITION, IT DISTRACTS THE RESEARCHER AND/OR HIS TECHNICIANS FROM LABORATORY WORK AND PERHAPS FROM A MORE COMPREHENSIVE ANALYSIS OF THE DATA. AN EASY TO USE AUTOMATED SYSTEM WOULD BE OF GREAT BENEFIT TO RESEARCHERS, PROVIDED IT WAS IN PRACTICE EASY TO USE. THIS REQUIREMENT IS ONE OF THE MOST IMPORTANT SPECIFICATIONS OF SUCH A SYSTEM. SINCE MOST BIOMEDICAL RESEARCHERS ARE NOT READILY CONVERSANT WITH COMPUTER SYSTEMS, NOR DO THEY GENERALLY HAVE THE TIME TO DEVOTE TOWARDS BECOMING SO. IN THEORY, ALL A RESEARCHER SHOULD HAVE TO DO IS TO WRITE HIS DATA ON A STANDARD FORM IN AN ORGANIZED MANNER, SPECIFY THE STATISTICS HE DESIRES, AND SUBMIT HIS PROBLEM TO A CENTRAL FACILITY. ALL HIS COMPUTATIONS WOULD THEN BE HANDLED BY AN AUTOMATED SYSTEM.

A **GENERALIZED RESEARCH ANALYSIS STATISTICAL SYSTEM, GRASS**, HAS BEEN WRITTEN BY THE DEPARTMENT OF INFORMATION SCIENCES TO PROVIDE THIS SUPPORT TO LABORATORY INVESTIGATORS. **GRASS** HAS THE CAPABILITY TO COMPUTE A VARIETY OF DESCRIPTIVE STATISTICS, TO PERFORM MANY PARAMETRIC AND NON-PARAMETRIC STATISTICAL TESTS, AND TO PRODUCE PLOTS AND HISTOGRAMS. THE DESCRIPTIVE STATISTICS INCLUDE MEANS, MEDIANS, STANDARD DEVIATIONS, STANDARD ERRORS OF THE MEAN, MINIMA, MAXIMA, RANGES, PEARSON CORRELATION COEFFICIENTS AND KENDALL†S RANK CORRELATION COEFFICIENTS. THE STATISTICAL TESTS INCLUDE PAIRED AND NON-PAIRED T-TESTS, WILCOXON†S RANK SUM AND SIGNED RANKS TESTS, GENERAL ONE-WAY ANALYSIS OF VARIANCE, DUNCAN†S NEW MULTIPLE RANGE TEST, NEWMAN-KEUL†S RANGE TEST, KRUSKAL-WALLIS TEST, MULTIPLE COMPARISONS BASED ON THE KRUSKAL-WALLIS STATISTIC, CHI-SQUARE TESTS, TESTS OF SIGNIFICANCE OF KENDALL†S CORRELATION COEFFICIENT, TESTS OF SIGNIFICANCE OF SIMPLE LINEAR, MULTIPLE, STEP-WISE, AND/OR POLYNOMIAL REGRESSIONS, AND NORMALITY TESTS.

GRASS ALSO HAS A VARIETY OF DATA TRANSFORMATION AND DATA MANIPULATION CAPABILITIES (I.E., NEW VARIABLES CAN BE CREATED AS FUNCTIONS OF EXISTING VARIABLES OR EXISTING VARIABLES CAN BE REWRITTEN).

ALL THAT IS REQUIRED OF THE INVESTIGATOR IS TO PREPARE A PUNCHED CARD DECK WHICH CONSISTS OF THE EXPERIMENTAL DATA TO BE ANALYZED AND A LIST OF THE ANALYSES TO BE PERFORMED. FOR USER CONVENIENCE, **GRASS** WILL ACCEPT THE EXPERIMENTAL DATA FROM INPUT SOURCES OTHER THAN CARDS. **GRASS** ALSO ALLOWS THE USER TO WRITE VARIABLES OF HIS CHOICE TO A VARIETY OF OUTPUT STORAGE FACILITIES.

PREPARATION OF THE PUNCHED CARD DECK

PREPARATION OF THE EXPERIMENTAL DATA FOR ANALYSIS BY THE **GRASS** PROGRAM REQUIRES THAT A DECK OF PUNCHED CARDS BE ASSEMBLED. A PUNCHED CARD IS A SPECIAL PURPOSE CARD WHICH CAN BE READ BY A DIGITAL COMPUTER SYSTEM. IT HAS 80 SEQUENTIALLY NUMBERED CONTIGUOUS COLUMNS. DATA WITH ONE CHARACTER PER COLUMN IS PUNCHED ONTO THESE CARDS. INPUT DATA ARE PUNCHED EITHER IN FIXED OR FORMAT-FREE DATA FIELDS. A DATA FIELD IS ONE OR MORE CONTIGUOUS COLUMNS ON A SINGLE CARD WHICH ARE ASSIGNED TO A VARIABLE.

THE **GRASS** DECK IS ASSEMBLED USING THREE TYPES OF CARDS - SYSTEM CONTROL CARDS, **GRASS** CONTROL CARDS, AND EXPERIMENTAL DATA CARDS. THESE CARDS MUST BE IN A SPECIFIC SEQUENCE FOR THE **GRASS** PROGRAM TO WORK PROPERLY. THE SYSTEM CONTROL CARDS ARE REQUIRED BY ALL JOBS.

THE SYSTEM CONTROL CARDS FOR THE **CDC 7600 BKY** OPERATING SYSTEM, (MODIFIED **SCOPE**), NECESSARY FOR PROCESSING A **GRASS** RUN WHEN EXPERIMENTAL DATA IS INPUT VIA CARDS ARE LISTED IN APPENDIX A. APPENDIX A ALSO CONTAINS EXAMPLES OF THE SYSTEM CARDS TO USE WHEN THE INPUT DEVICE IS OTHER THAN CARDS. SINCE THE EXAMPLES GIVEN APPLY ONLY TO THE **CDC 7600 BKY** OPERATING SYSTEM, THE USER SHOULD CONSULT A LOCAL COMPUTER SPECIALIST CONCERNING THE APPROPRIATE SYSTEM CARDS TO USE FOR A PARTICULAR MACHINE.

1.2 PREPARATION OF THE EXPERIMENTAL DATA CARDS

THE FINAL FORM OF THE INPUT DATA MAY BE REPRESENTED SCHEMATICALLY AS

```
                   VARIABLES
            X(1)   X(2)   X(3)..........X(N)
  CASES  *-----------------------------------*
    1    *  X11    X12    X13..........X1N   *
         *                                   *
    2    *  X21    X22    X23..........X2N   *
         *                                   *
    3    *  X31    X32    X33..........X3N   *
    .    *   .      .      .              .  *
    .    *   .      .      .              .  *
    .    *   .      .      .              .  *
    .    *   .      .      .              .  *
    M    *  XM1    XM2    XM3..........XMN   *
         *-----------------------------------*
```

WHERE X11 THROUGH X1N ARE ALL THE VALUES FOR THE FIRST CASE, (SUBJECT), AND THERE ARE M CASES. A CASE MAY BE PUNCHED ON AS MANY CARDS AS NECESSARY, BUT FOR FIXED FORMAT DATA A VALUE FOR A PARTICULAR VARIABLE MUST BE IN THE SAME LOCATION FROM CASE TO CASE. THAT IS IF EACH CASE REQUIRED 2 CARDS AND THE FIRST CASE†S WEIGHT WAS IN COLUMNS 1-4 OF THE FIRST CARD AND THE FIRST CASE†S HEIGHT WAS IN COLUMNS 10-13 OF THE SECOND CARD, ALL SUBSEQUENT CASES MUST HAVE WEIGHT PUNCHED IN COLUMNS 1-4 OF A CASE†S FIRST CARD AND HEIGHT PUNCHED IN COLUMNS 10-13 OF A CASE†S SECOND CARD. FOR FORMAT-FREE INPUT DATA, VARIABLES MUST BE IN THE SAME ORDER FROM CASE TO CASE.

THE REST OF THE DISCUSSION IN THIS SECTION APPLIES TO DATA SET UP IN FIXED FORMAT. FOR SETTING UP FORMAT FREE DATA SEE THE DISCUSSION OF THE $FMFREE CARD IN APPENDIX B.

TO PREPARE THE EXPERIMENTAL DATA CARDS IN THE GENERAL FORM ABOVE, SIX STEPS ARE NECESSARY. THESE STEPS MUST BE FOLLOWED CAREFULLY IF THE DESIRED RESULTS ARE TO BE ATTAINED. THE STEPS ARE?

1. SEQUENTIALLY NUMBER EACH VARIABLE.

2. DETERMINE THE MAXIMUM FIELD WIDTH, NUMBER OF CHARACTERS OF EACH VARIABLE. THIS WILL BE THE MINIMUM DATA FIELD SPECIFICATION.

3. LAYOUT THE DATA FIELDS FOR EACH VARIABLE ON STANDARD 80 COLUMN DATA SHEETS.

4. DETERMINE IF THE DECIMAL POINT IS TO BE PUNCHED OR LOCATE THE POSITION OF THE DECIMAL IN THE DATA FIELD IF THE DECIMAL POSITION IS TO BE IMPLIED THROUGH THE FORMAT STATEMENT.

5. TABULATE THE RAW DATA IN THE APPROPRIATE DATA FIELD ON THE 80 COLUMN DATA SHEETS.

6. DESCRIBE THE LAYOUT OF THE DATA FIELDS IN A FORMAT STATEMENT.

1. SEQUENTIALLY NUMBER EACH VARIABLE.

EACH OF THE VARIABLES IN THE STUDY MUST BE NUMBERED SEQUENTIALLY. THE VARIABLE NUMBER WILL BE USED TO IDENTIFY THE VARIABLE IN THE **GRASS** PROGRAM.

2. DETERMINE THE MINIMUM FIELD WIDTH FOR EACH VARIABLE.

THE TOTAL NUMBER OF CHARACTERS NECESSARY FOR A NUMERICAL OBSERVATION INCLUDES THE DIGITS, THE DECIMAL POINT, THE ALGEBRAIC SIGN, AND WHEN NECESSARY AN INDICATOR WHICH SPECIFIES THE EXPONENT OF BASE TEN BY WHICH THE DECIMAL DIGITS ARE TO BE MULTIPLIED. IF NO SIGN IS PUNCHED THE SYSTEM WILL ASSUME THE NUMBER IS POSITIVE. THE DECIMAL POINT MAY BE IMPLIED THROUGH THE USE OF A FORMAT STATEMENT, AS WILL BE DESCRIBED LATER. THE MINIMUM FIELD WIDTH MUST ALLOW FOR ALL CHARACTERS WHICH ARE NEEDED TO DESCRIBE A VARIABLE FOR ALL THE CASES. A FIELD WIDTH GREATER THAN THE MINIMUM MAY BE USED. UNUSED CHARACTERS MUST BE BLANKS OR ZEROES. USUALLY BLANKS PRECEDE THE FIRST CHARACTER OF A NUMBER AND ZEROES ARE USED AFTER THE FIRST CHARACTER HAS BEEN INDICATED. AN ENTIRE FIELD WHICH IS LEFT BLANK IS CONSIDERED A MISSING DATA POINT BY THE **GRASS** SYSTEM. A FIELD IN WHICH THE VALUE OF A VARIABLE IS ZERO MUST HAVE A ZERO PUNCHED IN IT.

3. LAYOUT THE DATA FIELD FOR EACH VARIABLE ON 80 COLUMN DATA SHEETS.

LAIR USERS MAY OBTAIN STANDARD 80 COLUMN DATA SHEETS FROM THE DEPARTMENT OF INFORMATION SCIENCES. OTHER USERS SHOULD CONTACT THEIR LOCAL COMPUTER FACILITY TO OBTAIN THESE FORMS. EACH COLUMN ON A SHEET CORRESPONDS TO A COLUMN ON AN 80 COLUMN HOLLERITH CARD. A KEYPUNCH OPERATOR WORKS DIRECTLY FROM THE CODING SHEETS IN PREPARING THE PUNCHED CARD DECK. THE DATA SHEET IS DIVIDED INTO DATA FIELDS WITH ONE FIELD FOR EACH VARIABLE. DATA FIELDS MAY BE SEPARATED BY BLANK FIELDS. FOR CLARIFICATION, DATA FIELDS MAY BE EMPHASIZED BY DRAWING HEAVY VERTICAL LINES ON THE SHEETS TO SEPARATE THE FIELDS. ENOUGH SHEETS ARE LAID OUT SO THAT EACH VARIABLE FOR A SINGLE CASE IS ASSIGNED AN ADEQUATELY SIZED DATA FIELD.

4. DETERMINE IF THE DECIMAL POINT IS TO BE PUNCHED OR LOCATE THE POSITION OF THE DECIMAL IN THE DATA FIELD IF IT IS IMPLIED IN THE FORMAT STATEMENT.

PUNCHING OF A DECIMAL POINT IS OPTIONAL BECAUSE A FORMAT STATEMENT DESCRIBES EACH DATA FIELD IN DETAIL INCLUDING THE LOCATION OF THE DECIMAL POINT. THE NUMBER 3.1416 COULD BE ENTERED ON THE DATA SHEET AS 31416 AND COULD BE READ BY THE **GRASS** PROGRAM AS 3.1416 IF SO SPECIFIED IN THE FORMAT STATEMENT. WHEN DECIMAL POINTS ARE NOT PUNCHED, DATA MUST BE ALIGNED IN THE DATA FEILD SO THAT THE FIRST DIGIT FOLLOWING THE IMPLIED DECIMAL POSITION IS ALWAYS ENTERED IN THE SAME COLUMN OF THE FIELD.

5. TABULATE THE RAW DATA IN APPROPRIATE FIELDS.

ENTER ALL DATA FOR ONE CASE ON A HORIZONTAL LINE ON ONE OR MORE DATA SHEETS WHICH HAVE BEEN PREPARED. SUCCESSIVE CASES ARE ENTERED ON SUBSEQUENT LINES IN TABULAR FASHION. WHEN A SINGLE SET OF DATA SHEETS ARE EXHAUSTED, NEW SETS OF SHEETS ARE USED WITH THE EXACT FIELD SPECIFICATIONS OF THE FIRST SET.

6. DESCRIBE THE LAYOUT OF THE DATA FIELDS IN A FORMAT STATEMENT

THE FORMAT STATEMENT IS A DESCRIPTION OF THE WAY IN WHICH THE DATA SHEETS ARE LAID OUT AND CONSEQUENTLY, THE WAY IN WHICH THE DATA WILL APPEAR ON PUNCHED CARDS. THE **GRASS** FORMAT STATEMENT BEGINS WITH A DOLLAR SIGN, **$**, IN COLUMN 1, FOLLOWED BY THE WORD **FORMAT** IN COLUMNS 2-7. FOLLOWING THIS IS A STANDARD **FORTRAN IV** FORMAT STATEMENT. THE SPECIFICATIONS WHICH MAY BE USED ARE **F, E,** AND/OR **X.** NUMERICAL DATA MAY BE OF EITHER F OR **E** TYPE. **F** OR **E** TYPE DATA ARE READ USING **F** OR **E** TYPE SPECIFICATIONS RESPECTIVELY. BOTH TYPES OF DATA CONSIST OF ONE OR MORE DIGITS WITH A SINGLE DECIMAL POINT AND A SIGN. AN **E** TYPE NUMBER DIFFERS FROM AN **F** TYPE NUMBER IN THAT IT ALSO HAS AN EXPONENT FOLLOWING THE DIGITS. THE EXPONENT IS THE LETTER -E- FOLLOWED BY AN OPTIONAL SIGN POSITION FOLLOWED BY ONE OR TWO DIGITS. THE MEANING OF THE EXPONENT IS THAT THE PRECEDING NUMBER IS TO BE MULTIPLIED BY TEN RAISED TO THE POWER GIVEN BY THE SIGN AND THE FOLLOWING DIGITS, (I.E. 2.5E10=2.5 TIMES 10 TO THE 10TH POWER). SOME EXAMPLES OF **F** AND **E** TYPE NUMBERS AND THE NUMERICAL VALUES THEY REPRESENT ARE

NUMERIC VALUE	E TYPE	F TYPE
200	2.00E02	200.0
-10.5	-1.05E01	-10.5
.0033	3.3E-03	.0033

THE GENERAL FORM OF AN **F** OR **E** TYPE SPECIFICATION IS **NCW.D** WHERE

- N - NUMBER OF SUCCESSIVE REPETITIONS OF THIS FIELD
- C - TYPE OF SPECIFICATION, F OR E
- W - AN INTEGER SPECIFYING THE WIDTH OF THE FIELD INCLUDING DECIMAL POINT (IF ANY) AND SIGN (IF ANY)
- D - NUMBER OF DECIMAL PLACES UNDERSTOOD TO BE IN THE FIELD.

SOME EXAMPLES ARE

NUMERIC VALUE	INPUT DATA	SPECIFICATION
22.5	22.5	F4.1
999.9	9999	F4.1
99.99	9999	F4.2
9.999	9999	F4.3
-.00001	-1.0E-05	E8.1
155.7	1.557E02	E8.3
1000	10.00E02	E8.2

IT CAN BE SEEN IN THE ABOVE THAT -W-, THE WIDTH OF THE FIELD INCLUDES

THE DECIMAL POINT AND SIGN IF THESE ARE PRESENT. FURTHER BY CHOOSING AN APPROPRIATE VALUE FOR -D-, THE DECIMAL POINT CAN BE IMPLIED WHEN USING F TYPE NUMBERS. A DECIMAL POINT IS USUALLY PUNCHED WHEN THE RAW DATA IS OF THE E TYPE. A USEFUL RULE TO REMEMBER IS THAT A PUNCHED DECIMAL POINT OVERRIDES THE -D- CHARACTER SPECIFIED IN THE FORMAT. THUS IF THE NUMBER **22.5** IS PUNCHED, ANY OF THE FORMATS **F4.4, F4.3, F4.2, F4.1,** OR **F4.0** WOULD CAUSE THE VALUE **22.5** TO BE READ.

THE **X** SPECIFICATION IS USED TO SKIP OVER BLANKS OR NUMERIC DATA. THE GENERAL FORM OF THE **X** SPECIFICATION IS NX WHERE N REPRESENTS THE NUMBER OF COLUMNS TO BE SKIPPED. THUS 10X WOULD CAUSE 10 COLUMNS TO BE SKIPPED.

ELEMENTS OF A FORMAT STATEMENT ARE SEPARATED BY COMMAS. IF THE VARIABLES FOR A CASE ARE PUNCHED ON MORE THAN ONE CARD, A SLASH, (/), INDICATES THAT A NEW CARD IS TO BE READ. SINCE THE FORMATS FOR ALL CASES ARE IDENTICAL, ONLY ONE $FORMAT CARD IS USED. THE **GRASS** PROGRAM WILL READ AS MANY CASES AS ARE SPECIFIED ON THE $DATA CARD.

EXAMPLES OF GRASS FORMAT STATEMENTS

1. IN A HYPOTHETICAL STUDY, 4 VARIABLES ARE MEASURED. THE VALUES FOR THE FIRST CASE ARE **35.5, 799, 81000,** AND **2.** THESE MIGHT BE PUNCHED ON CARDS AS

```
                           1 1 1 1 1 1 1 1 1 1 2    8
COLUMN   1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 -- 0

VALUES       3 5 5     7 9 9 8 1 0 0 0         2
```

USING THE **F** AND **X** SPECIFICATIONS THIS DATA COULD BE READ USING THE FORMAT STATEMENT

$FORMAT(2X,F3.1,2X,F3.0,F5.0,4X,F1.0).

IF WE CONSIDER HOWEVER THAT ALL NUMBERS ARE IN FIELDS OF 5, (THE FIRST AND SECOND NUMBER HAVE 2 LEADING BLANKS, AND THE FOURTH NUMBER HAS 4 LEADING BLANKS), WE SEE THAT THE INPUT DATA COULD BE READ WITH THE FORMAT STATEMENT

$FORMAT(F5.1,F5.0,F5.0,F5.0).

THIS FORMAT STATEMENT COULD BE FURTHER CONTRACTED. NOTE THAT THE LAST THREE ELEMENTS ARE **F5.0.** THESE CAN BE COMBINED AS **3F5.0** (THREE SUCCESIVE FIELDS IN F5.0 FORMAT ARE TO BE READ). THUS THE FORMAT STATEMENT COULD BE REWRITTEN AS

$FORMAT(F5.1,3F5.0).

THIS IS A MUCH SHORTER FORM THAN THE FIRST FORMAT STATEMENT BUT CONTAINS THE SAME INFORMATION AND WILL RESULT IN THE SAME NUMBERS, **35.5, 799, 81000** AND **2** BEING READ.

2. A HYPOTHETICAL STUDY INVOLVES THE MEASUREMENT OF 5 VARIABLES. THE VALUES FOR THE FIRST CASE ARE **.211, .333, 2100, 57.8,** AND **100.**

PREPARATION OF THE EXPERIMENTAL DATA CARDS

THESE VALUES MIGHT BE ARRANGED AND PUNCHED AS

```
                                   1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2      8
   COLUMN  1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 ---  0

             . 2 1 1   . 3 3 3   2 1 0 0     5 7 8     1 0 0
```

THE READER SHOULD CONVINCE HIMSELF THAT ANY OF THE FOLLOWING FORMAT STATEMENTS WILL RESULT IN THE VALUES .211, .333, 2100, 57.8, AND 100 BEING READ.

```
$FORMAT(1X,F4.3,1X,F4.3,1X,F4.0,2X,F3.1,2X,F3.0)
$FORMAT(2(1X,F4.3),1X,F4.0,2X,F3.1,2X,F3.0)
$FORMAT(F5.3,F5.3,F5.0,F5.1,F5.0)
$FORMAT(2F5.3,F5.0,F5.1,F5.0)
$FORMAT(F5.0,F5.0,F5.0,F5.1,F5.0)
$FORMAT(3F5.0,F5.1,F5.0)
```

1.3 GRASS CONTROL CARDS

OVERVIEW

THERE ARE THREE TYPES OF GRASS CONTROL CARDS - INFORMATION CONTROL CARDS, STATISTICAL ROUTINE SPECIFICATION CARDS, AND DATA TRANSFORMATION CARDS. THE INFORMATION CONTROL CARDS SPECIFY GENERAL INFORMATION TO THE GRASS PROGRAM, SUCH AS THE NUMBER OF VARIABLES AND OBSERVATIONS AND THE FORMAT OF THE INPUT DATA. THE STATISTICAL ROUTINE SPECIFICATION CARDS SPECIFY WHICH STATISTICAL ROUTINES ARE TO BE USED IN THE ANALYSES. THE DATA TRANSFORMATION CARDS SPECIFY THE TYPES OF DATA TRANSFORMATIONS AND MANIPULATIONS THE USER DESIRES TO PERFORM ON THE DATA.

ALL GRASS CONTROL CARDS HAVE A DOLLAR SIGN, $, PUNCHED IN COLUMN 1, FOLLOWED BY A GRASS CONTROL WORD PUNCHED IN COLUMNS 2 THROUGH 7, FOLLOWED BY A PARAMETER FIELD PUNCHED IN COLUMNS 8 THROUGH 79.

*NOTE THE FOLLOWING IN THE PREPARATION OF INFORMATION CONTROL CARDS AND STATISTICAL ROUTINE SPECIFICATION CARDS-

1) IF THE PARAMETER FIELD HAS TO BE CONTINUED ON SUBSEQUENT CARDS, THE LETTER -C- MUST BE PUNCHED IN COLUMN 80 OF THE CURRENT CARD.

2) WHEN A CONTINUATION CARD IS NECESSARY THE CARD LABEL, CONTROL WORD, **MUST NOT** BE INCLUDED ON THE CONTINUATION CARD.

3) THE PARAMETER FIELDS OF CERTAIN CARDS MAY NOT BE CONTINUED ONTO SUBSEQUENT CARDS. THIS LIMITATION IS NOTED FOR AFFECTED CONTROL CARDS IN APPENDICES B AND C.

INFORMATION CONTROL CARDS

THE FOLLOWING IS AN INDEX OF GRASS INFORMATION CONTROL CARDS. THESE CARDS ARE DESCRIBED IN DETAIL IN APPENDIX B.

$TITLE A TITLE OF THE USER'S CHOICE IS PUNCHED IN THE PARAMETER FIELD. THIS TITLE APPEARS AT THE TOP OF EACH PAGE OF OUTPUT, UNTIL ANOTHER $TITLE CARD IS USED.

$DATA THE PARAMETER FIELD INCLUDES THE NUMBER OF VARIABLES, NUMBER OF OBSERVATIONS, AND AN INDICATION OF WHETHER THE INPUT DATA IS TO BE PRINTED IN THE OUTPUT.

$FORMAT THE PARAMETER FIELD IS A FORTRAN FORMAT STATEMENT IN WHICH F, E, OR X TYPE SPECIFICATIONS MAY BE USED TO READ INPUT DATA WHICH ARE PUNCHED OR WRITTEN IN FIXED FORMAT.

$FMFREE THERE IS NO PARAMETER FIELD. DATA PUNCHED IN A FORMAT-FREE MANNER, (EACH VALUE HAS THE DECIMAL POINT PUNCHED AND VALUES ARE SEPARATED BY BLANK OR COMMAS), WILL BE READ.

$WRITE THE PARAMETER FIELD INCLUDES A SPECIFICATION OF WHICH VARIABLES ARE TO BE WRITTEN.

$FILE THE PARAMETER FIELD IS A 1 THROUGH 7 CHARACTER NAME. IF $FILE IS USED AFTER THE $DATA CARD, THE FORMATTED INPUT DATA WILL BE READ FROM THE FILE NAMED IN THE PARAMETER FIELD. IF THE $FILE CARD IS USED AFTER THE $WRITE CARD, THE VARIABLES SPECIFIED ON THE $WRITE CARD WILL BE WRITTEN TO THE FILE NAMED OR TO BCDOUT IF THE PARAMETER FIELD ON THE $FILE CARD IS LEFT BLANK. A $FORMAT CARD MUST FOLLOW THE $FILE CARD. THE $FILE CARD IS NOT USED WHEN INPUT DATA IS ON CARDS.

$FILEB THE PARAMETER FIELD INCLUDES A 1 THROUGH 7 CHARACTER NAME. IF USED AFTER THE $DATA CARD, THE NON-FORMATTED INPUT DATA WILL BE READ FROM THE FILE NAMED. IF USED AFTER THE $WRITE CARD, THE VARIABLES SPECIFIED ON THE $WRITE CARD WILL BE WRITTEN IN NON-FORMATTED, (BINARY), FORM TO THE FILE NAMED. NO $FORMAT CARD SHOULD FOLLOW THE $FILEB CARD. THE $FILEB CARD IS NOT USED WHEN THE INPUT DATA IS ON CARDS IN NON-FORMATTED FORM.

$CMT THE PARAMETER FIELD INCLUDES ANY COMMENT THE USER WISHES TO MAKE. THE COMMENT(S) WILL BE PRINTED IN THE OUTPUT.

$RWD THE PARAMETER FIELD INCLUDES A 1 THROUGH 7 CHARACTER NAME. THE FILE NAMED WILL BE REWOUND.

$SCALE THE PARAMETER FIELD SPECIFIES THE SCALE TO BE USED IN SETTING UP THE AXES FOR A PLOT OR HISTOGRAM.

$ADD THE PARAMETER FIELD SPECIFIES TWO VARIABLES WHICH ARE TO BE COMBINED AND A THIRD VARIABLE WHICH WILL CONTAIN THE COMBINED VARIABLES.

$END THIS IS THE LAST GRASS CONTROL CARD OF EVERY DECK.

$TITLE,$DATA AND $END ARE REQUIRED ON ALL GRASS RUNS. CONDITIONAL CARDS ARE AS FOLLOWS--

1) $FORMAT MUST BE USED WHEN INPUT DATA IS ON CARDS IN FORMATTED FORM OR WHEN THE $FILE CARD IS USED.

2) $FILE MUST BE USED WHEN INPUT DATA IS IN FORMATTED FORM ON A SOURCE OTHER THAN CARDS.

3) $FMFREE MUST BE USED WHEN INPUT DATA IS ON CARDS IN NON-FORMATTED FORM.

4) $FILEB MUST BE USED WHEN INPUT DATA IS IN NON-FORMATTED, BINARY, FORM ON A SOURCE OTHER THAN CARDS. ALL OTHER INFORMATION CONTROL CARDS ARE OPTIONAL.

APPENDICES A AND D CONTAIN SOME SAMPLE JOB SETUPS.

STATISTICAL ROUTINE SPECIFICATION CONTROL CARDS

THE FOLLOWING IS AN INDEX OF THE STATISTICAL ROUTINE SPECIFICATION CONTROL CARDS. A THOROUGH DESCRIPTION OF THESE CARDS AND THE ANALYSES PERFORMED ARE PRESENTED IN APPENDIX C.

$BSTAT THE NUMBER OF OBSERVATIONS, MINIMUM, MAXIMUM, RANGE, MEDIAN, MEAN, STANDARD DEVIATION, AND THE STANDARD ERROR OF THE MEAN ARE PRINTED FOR EACH VARIABLE REQUESTED.

$NTST NON-PAIRED T-TESTS ARE PERFORMED BETWEEN PAIRWISE COMBINATIONS OF ALL VARIABLES SPECIFIED IN THE PARAMETER FIELD.

$PTST PAIRED T-TESTS ARE PERFORMED BETWEEN PAIRWISE COMBINATIONS OF ALL VARIABLES SPECIFIED.

$RNKSM WILCOXON'S RANK SUM STATISTIC IS CALCULATED FOR THE 2 VARIABLES SPECIFIED. THE LARGE SAMPLE APPROXIMATION OF THE RANK SUM STATISTIC IS ALSO PRINTED. IN ADDITION, THE SIEGEL-TUKEY TEST STATISTIC IS PRINTED.

$SNRNK WILCOXON'S SIGNED RANK STATISTIC AND ITS LARGE SAMPLE APPROXIMATION ARE COMPUTED FOR THE 2 VARIABLES SPECIFIED.

$ANOVA A ONE-WAY ANALYSIS OF VARIANCE FOR EQUAL OR UNEQUAL NUMBERS OF REPLICATES IS PERFORMED FOR THE VARIABLES REQUESTED.

$DNMRT DUNCAN'S NEW MULTIPLE-RANGE TEST IS USED TO COMPARE THE MEANS OF THE VARIABLES SPECIFIED.

$NKMRT NEWMAN-KEUL'S TEST IS USED TO COMPARE THE MEANS OF THE VARIABLES SPECIFIED.

$KRUWL THE KRUSKAL-WALLIS H STATISTIC FOR TESTING THE HYPOTHESIS OF NO DIFFERENCES AMONG K TREATMENTS IS CALCUALTED FOR THE K VARIABLES SPECIFIED.

$KRUWC THE OUTPUT INCLUDES THE KRUSKAL-WALLIS H STATISTIC AND ALSO INCLUDES A MULTIPLE COMPARISON TEST BETWEEN ALL PAIRWISE COMBINATIONS OF THE K VARIABLES SPECIFIED.

$CORR PEARSON PRODUCT MOMENT CORRELATION COEFFICIENTS ARE COMPUTED FOR ALL PAIRWISE COMBINATIONS OF VARIABLES REQUESTED.

$RNKCR KENDALL'S RANK CORRELATION COEFFICIENTS ARE COMPUTED FOR ALL PAIRWISE COMBINATIONS OF VARIABLES REQUESTED.

$CHISQ THE CHI-SQUARE STATISTIC AND CONTINGENCY COEFFICIENT ARE CALCULATED FOR A TWO-WAY CONTINGENCY TABLE.

$LSQR A LEAST-SQUARES REGRESSION ANALYSIS IS PERFORMED FOR A SPECIFIED DEPENDENT VARIABLE AND A SPECIFIED INDEPENDENT VARIABLE.

$MREG	A MULTIPLE REGRESSION ANALYSIS IS PERFORMED FOR A SPECIFIED DEPENDENT VARIABLE AND SPECIFIED INDEPENDENT VARIABLES.
$SWREG	A STEPWISE MULTIPLE REGRESSION ANALYSIS IS PERFORMED FOR A SPECIFIED DEPENDENT VARIABLE AND SPECIFIED INDEPENDENT VARIABLES.
$POLRG	A POLYNOMIAL REGRESSION ANALYSIS IS PERFORMED ON THE DEPENDENT AND INDEPENDENT VARIABLES REQUESTED. A POLYNOMIAL OF DEGREE FIVE OR LESS MAY BE FITTED.
$NORMT	VARIOUS TESTS OF NORMALITY ARE PERFORMED ON EACH VARIABLE REQUESTED. THESE INCLUDE A CHI-SQUARE GOODNESS OF FIT TEST, AND TESTS OF SKEWNESS AND KURTOSIS.
$PLOT	A PLOT IS ACCOMPLISHED BETWEEN 2 VARIABLES.
$PLOTN	A PLOT OF A SPECIFIED DEPENDENT VARIABLE VERSUS EACH SPECIFIED INDEPENDENT VARIABLE IS OUTPUT.
$PLOTI	A PLOT OF A SPECIFIED INDEPENDENT VARIABLE VERSUS EACH SPECIFIED DEPENDENT VARIABLE IS OUTPUT.
$HSTGM	A HISTOGRAM, FREQUENCY DISTRIBUTION, IS PRINTED FOR EACH VARIABLE REQUESTED.
$FDIST	THE LEVEL OF SIGNIFICANCE OF THE INPUT F-VALUE IS PRINTED.
$CDIST	THE LEVEL OF SIGNIFICANCE OF THE INPUT CHI-SQUARE VALUE IS PRINTED.
$TDIST	THE LEVEL OF SIGNIFICANCE OF THE INPUT T-VALUE IS PRINTED.
$RSDST	THE LEVEL OF SIGNIFICANCE OF EITHER THE INPUT WILCOXON'S RANK SUM STATISTIC OR THE INPUT WILCOXON'S SIGNED RANK STATISTIC IS PRINTED.

DATA TRANSFORMATION CARDS - THE $TRANS STATEMENT

THE $TRANS CARDS GIVE THE GRASS USER A GREAT DEAL OF FLEXIBILITY IN TRANSFORMING AND/OR EDITING THE EXPERIMENTAL DATA.

THE CONTROL STATEMENT BEGINS WITH A DOLLAR SIGN, $, IN COLUMN 1 FOLLOWED BY THE CONTROL WORD TRANS IN COLUMNS 2 THROUGH 6. THE PARAMETER FIELD IN COLUMNS 8-79 CONTAINS A TRANSFORMATION STATEMENT. THERE ARE FOUR TYPES OF TRANSFORMATION STATEMENTS -- THE VARIABLE STATEMENT, THE TRANSPOSE STATEMENT, THE ASSIGNMENT STATEMENT, AND THE CONDITIONAL STATEMENT.

1. THE VARIABLE STATEMENT ALLOWS THE USER TO ASSIGN A NAME TO THE INPUT VARIABLES. NORMALLY AN INPUT VARIABLE IS REPRESENTED BY THE SUBSCRIPTED VARIABLE X,(I.E. VARIABLE 1 IS X(1), VARIABLE 10 IS X(10)). THE VARIABLE STATEMENT ALLOWS THE USER TO CHANGE THE NAME,X, TO ANY NAME. THE NEW NAME IS NOT OUTPUT AS A LABEL. AFTER SUPPLYING A NEW NAME WITH THE VARIABLE STATEMENT EITHER THE NEW NAME OR X MAY BE USED IN SUBSEQUENT TRANSFORMATION STATEMENTS. THIS ALLOWS SOME DIFFERENTIATION OR CLARIFICATION IN THE SUBSEQUENT STATEMENTS.

EXAMPLES
$TRANS VARIABLE DATA

INPUT VARIABLE 1 COULD BE CALLED DATA(1). INPUT VARIABLE 10 COULD BE CALLED DATA(10).

$TRANS VARIABLE CURVE

INPUT VARIABLE 1 COULD BE CALLED CURVE(1). INPUT VARIABLE 10 COULD BE CALLED CURVE(10).

2. THE TRANSPOSE TRANSFORMATION STATEMENT ALLOWS THE USER TO INTERCHANGE THE ROWS AND COLUMNS OF THE INPUT DATA. THIS EFFECTIVELY INTERCHANGES OBSERVATIONS (ROWS) AND VARIABLES(COLUMNS).

EXAMPLE
$TRANS TRANSPOSE

IF THE DATA WAS INPUT WITH 3 VARIABLES AND 2 OBSERVATIONS PER VARIABLE - I.E.,

```
  2  3  4
  5  6  7
```

THE ABOVE $TRANS STATEMENT WOULD RESULT IN DATA WITH 2 VARIABLES AND 3 OBSERVATIONS PER VARIABLE - I.E.,

```
  2  5
  3  6
  4  7
```

3. THE ASSIGNMENT STATEMENT ALLOWS THE USER TO DEFINE NEW VARIABLES OR TO REDEFINE OLD INPUT VARIABLES. THE FORM OF THE STATEMENT MAY BE ANY LOGICALLY CONSISTENT COLLECTION OF OPERANDS, (VARIABLES OR CONSTANTS), OPERATIONS, AND FUNCTIONS EXPRESSING A

RELATION BETWEEN A DEPENDENT VARIABLE AND ONE OR MORE INDEPENDENT VARIABLES. THE OPERATIONS WHICH ARE ALLOWED ARE?

+	ADDITION
-	SUBTRACITON
*	MULTIPLICATION
/	DIVISION
↑	EXPONENTIATION

UNLESS OTHERWISE INDICATED BY PARENTHESES THE OPERATIONS ARE PERFORMED IN THE FOLLOWING ORDER?

1) EXPONENTIATION IS PERFORMED FIRST.
2) MULTIPLICATION AND DIVISION ARE PERFORMED NEXT.
3) ADDITION AND SUBTRACTION ARE PERFORMED LAST.

OPERATIONS OF THE SAME PRIORITY (I.E. MULTIPLICATION AND DIVISION) ARE PERFORMED IN ORDER OF OCCURENCE FROM LEFT TO RIGHT.

EXAMPLES

GRASS EXPRESSION	ALGEBRAIC EQUIVALENT
A*B-C	(A)(B)-C
A*(B-C)	A(B-C)
A*B/C*D	(A)(B)(D)/(C)
A*B/(C*D)	(A)(B)/(C)(D)
A↑B+C	(A RAISED TO THE POWER B)+C
A↑(B+C)	A RAISED TO THE POWER (B+C)

THE FOLLOWING FUNCTIONS ARE AVAILABLE IN THE TRANS COMPILER

FUNCTION NAME	FUNCTION PERFORMED
SIN	SINE
COS	COSINE
TAN	TANGENT
ASIN	ARC SINE
ACOS	ARC COSINE
ATAN	ARC TANGENT
LN	LOG BASE E
LOG	LOG BASE 10
EXP	EXPONENTIATION BASE E
ABS	ABSOLUTE VALUE
SQRT	SQUARE ROOT
PROBIT	PROBIT OF A NUMBER BETWEEN 0 AND 1
APROBIT	INVERSE OF PROBIT
LINE	WHEN USED AFTER $LSQR OR $POLRG COMPUTES NEW DEPENDENT VARIABLE VALUES USING THE LEAST SQUARE ESTIMATES
RANF	
LOGICAL	LOGICAL VALUE 1 OR 0

EXAMPLE

```
$TRANS X(2)=  3*LN(1+SIN(X(1)))
```

THE ABOVE $TRANS STATEMENT WOULD CAUSE A NEW SECOND VARIABLE TO BE CREATED EQUAL TO 3 TIMES THE NATURAL LOG OF THE QUANTITY 1 PLUS THE SINE OF VARIABLE 1.

IN CREATING NEW VARIABLES THE FOLLOWING SHOULD BE NOTED?

A) WHEN ANY OF THE VALUES OF INDEPENDENT VARIABLES ARE MISSING THE CORRESPONDING VALUES OF THE DEPENDENT VARIABLE WILL BE MISSING.

B) IF THE DEPENDENT VARIABLE IN AN ASSIGNMENT STATEMENT IS AN ALREADY EXISTING VARIABLE, THE OLD VALUES WILL BE REWRITTEN AS DICTATED BY THE ASSIGMENT STATEMENT.

4. THE CONDITIONAL STATEMENTS TEST A GIVEN COMPARISON AND TAKE ONE OF TWO BRANCHES, (ASSIGNMENT STATEMENTS), DEPENDING ON THE RESULT OF THE COMPARISON. THE GENERAL FORM OF THE CONDITIONAL STATEMENT IS

```
$TRANS  IF A  THEN B
$TRANS  ELSE C
```

WHERE -A- IS A COMPARISON STATEMENT, AND -B- AND -C- ARE ASSIGNMENT STATEMENTS. A COMPARISON STATEMENT IS COMPOSED OF TWO EXPRESSIONS SEPARATED BY A SIGN OF COMPARISON, A COMPARATOR. THE COMPARATORS AVAILABLE TO THE USER ARE?

=	EQUAL
≠	NOT EQUAL
<	LESS THAN
≤	LESS THAN OR EQUAL
>	GREATER THAN
≥	GREATER THAN OR EQUAL

IF THE COMPARISON -A- PROVES TRUE, THEN THE STATEMENT DEFINED AS THE OBJECT OF THE **THEN** CLAUSE -B- IS EXECUTED. OTHERWISE THE OBJECT OF THE **ELSE** CLAUSE -C- IS EXECUTED. THE **ELSE** CLAUSE IS OPTIONAL. IF IT IS DESIRED TO LEAVE UNCHANGED THE VALUE OF THE DEPENDENT VARIABLE WHEN -A- PROVES FALSE, THE **ELSE** CLAUSE NEED NOT BE INCLUDED.

THE OBJECTS OF THE **THEN** AND **ELSE** CLAUSES MAY BE A **DO** STATEMENT, IN WHICH CASE ALL OF THE ASSIGNMENT STATEMENTS FOLLOWING THE **DO** STATEMENT WILL BE EXECUTED UNTIL AN **END** STATEMENT TERMINATES THE LIST. **DO** STATEMENTS MAY ALSO BE NESTED. THIS FACILITY ALLOWS THE EXECUTION OF SEVERAL ASSIGNMENT STATEMENTS UPON THE RESULT OF A SINGLE COMPARISON STATEMENT.

IN ADDITION, THE OBJECTS OF THE **THEN** OR **ELSE** STATEMENT MAY BE ANOTHER **IF** STATEMENT ALLOWING FURTHER FLEXIBILITY.

EXAMPLES

```
1. $TRANS  IF X(2)<LOG(X(1)) THEN X(3)=5*X(1)
   $TRANS  ELSE X(3)=1-4*ATAN(X(3))
```

WHENEVER A VALUE OF X(2) IS LESS THAN LOG BASE 10 OF X(1) THE THIRD VARIABLE, X(3), WOULD BE ASSIGNED A VALUE EQUAL TO 5 TIMES THE VALUE OF X(1), OTHERWISE THE THIRD VARIABLE WOULD BE ASSIGNED A VALUE EQUAL TO 1 MINUS 4 TIMES THE ARCTANGENT OF THE THIRD VARIABLE. THIS WOULD NATURALLY RESULT IN REPLACEMENT OF VALUES PREVIOUSLY ASSIGNED TO THE THIRD VARIABLE.

```
2. $TRANS  IF X(2)=0  THEN IF LOGICAL(X(2))=1 THEN X(2)=1
```

THIS STATEMENT WOULD DETECT MISSING VALUES IN THE SECOND VARIABLE, X(2), AND ASSIGN A VALUE OF 1 WHERE THE MISSING VALUES OCCURED.

```
3. $TRANS  IF X(3)>0 THEN DO
   $TRANS  X(1)=X(3)*5-2
   $TRANS  X(2)=SQRT(X(4))
   $TRANS  END
   $TRANS  ELSE DO
   $TRANS  X(1)=0
   $TRANS  X(2)=0
   $TRANS  X(3)=1
   $TRANS  END
```

WHENEVER A VALUE OF THE THIRD VARIABLE, X(3), IS GREATER THAN 0, THE FIRST VARIABLE, X(1), WILL BE ASSIGNED A VALUE EQUAL TO (X(3)(5))-2 AND THE SECOND VARIABLE WILL BE ASSIGNED A NEW VALUE EQUAL TO THE SQUARE ROOT OF THE FOURTH VARIABLE, OTHERWISE, X(1) AND X(2) WILL BE ASSIGNED A VALUE OF 0 AND X(3) WILL BE REPLACED BY THE VALUE 1.

***NOTE** THE FOLLOWING REGARDING THE USE OF **$TRANS** CARDS.

1) FOLLOWING ANY TRANSFORMATION THE NUMBER OF VARIABLES AND OBSERVATIONS IS AUTOMATICALLY UPDATED. THE $DATA CARD SHOULD REFLECT THE NUMBER OF VARIABLES AND OBSERVATIONS INPUT AND **NOT** THE NUMBERS FOLLOWING TRANSFORMATIONS.

2) GROUPS OF **$TRANS** CARDS WHICH APPEAR AFTER THE INPUT DATA HAS BEEN ENTERED MUST BE FOLLOWED BY A **$EXECU** CARD OR ANY STATISTICAL ROUTINE SPECIFICATION CARD. FOR EXAMPLE, IF THE CARDS IN EXAMPLE 3 WERE INSERTED AFTER THE INPUT DATA HAD BEEN ENTERED, A **$EXECU** CARD OR A STATISTICAL ROUTINE SPECIFICATION CARD MUST FOLLOW THE **$TRANS END** STATMENT BEFORE ADDITIONAL **$TRANS** STATEMENTS COULD BE INSERTED.

3) NO MORE THAN 10,000 DATA ELEMENTS MAY BE HANDLED WITH ONE SET OF **$TRANS** CARDS (I.E., ONE ASSIGNEMNT OR TRANSPOSE STATEMENT OR A CONDITIONAL STATEMENT WITH ITS ASSIGNMENT STATEMENTS).

1.4 SAMPLE PROBLEMS

FIVE SAMPLE PROBLEMS ARE PRESENTED TO SHOW ASSEMBLED CARD DECKS AND OUTPUT FOR SOME TYPICAL GRASS RUNS. IN THE FIRST EXAMPLE, THE SIX STEPS USED IN PREPARING THE EXPERIMENTAL DATA WILL BE DETAILED. IN ALL EXAMPLES, THE DATA ARE FICTITIOUS.

EXAMPLE 1

AN INVESTIGATOR HAS 4 VARIABLES WITH 6 OBSERVATIONS OF EACH VARIABLE. HE DESIRES TO COMPUTE MEANS AND STANDARD DEVIATIONS FOR EACH VARIABLE, CORRELATIONS BETWEEN TIME ON TREADMILL, AGE, BODY TEMPERATURE AND PULSE, AND A REGRESSION LINE RELATING TIME OF WALK TO THE OTHER 3 VARIABLES. THE RAW DATA ARE SHOWN IN TABLE 1.

TABLE 1

SUBJECT NUMBER	TREADMILL SPEED(MPH)	AGE	TIME OF WALK(MIN)	BODY TEMP.	PULSE (BPM)
1	5	20	10.0	98.9	123
2	5	18	15.8	98.7	147
3	5	35	8.8	99.1	109
4	5	24	8.0	98.6	118
5	5	41	12.5	99.4	131
6	5	43	12.8	99.7	135

THE STEPS FOR PUTTING THIS DATA IN A SUITABLE FORM FOR GRASS USE ARE AS FOLLOWS.

1. ASSIGN SEQUENTIAL NUMBERICAL IDENTIFIERS TO EACH VARIABLE.

VARIABLE	NUMERIC IDENTIFIER
AGE	1
TIME OF WALK	2
BODY TEMP.	3
PULSE	4

2. DETERMINE THE MAXIMUM FIELD WIDTH (NUMBER OF CHARACTERS) FOR EACH VARIABLE.

VARIABLE	MAXIMUM FIELD WIDTH
AGE	2
TIME OF WALK	4 (3)
BODY TEMP.	4 (3)
PULSE	3

NUMBERS IN PARENTHESES ARE THE MAXIMUM FIELD WIDTH IF THE DECIMAL IS NOT PUNCHED.

3,4,5. THE DATA ARE ENTERED ON 80 COLUMN DATA SHEETS, (FIGURE 1), FOR KEYPUNCHING. IN THIS EXAMPLE, THE DECIMAL POINT IS TO BE PUNCHED FOR TIME OF WALK AND WILL BE IMPLIED FOR BODY TEMPERATURE.

6. THE DATA COULD BE DESCRIBED BY THE FORMAT STATEMENT, **$FORMAT(F2.0,F5.0,F4.1,F4.0).**

THE ASSEMBLED CARD DECK, INCLUDING SYSTEM CONTROL CARDS, INFORMATION CONTROL CARDS, DATA CARDS AND STATISTICAL ROUTINE SPECIFICATION CARDS, IS SHOWN IN FIGURE 2. NOTE THAT THE DATA CARDS IMMEDIATELY FOLLOW THE **$FORMAT** CARD.

COMPUTER OUTPUT FOR THIS RUN IS GIVEN IN FIGURES 3 THROUGH 15. FIGURE 4 IS A LISTING OF THE RAW DATA. FIGURE 5 IS THE OUTPUT FROM THE BASIC STATISTICS ROUTINE. FIGURE 6 SHOWS THE PEARSON CORRELATION COEFFICIENTS FOR ALL VARIABLES. FIGURES 7 THROUGH 12 GIVE THE OUTPUT FROM THE RANK CORRELATION ROUTINE. FIGURES 13 THROUGH 15 ARE THE RESULTS FROM MULTIPLE REGRESSION.

EXAMPLE 2

AN INVESTIGATOR HAS COLLECTED DATA ON 8 SUBJECTS. THE VARIABLES MEASURED ARE **HEIGHT, WEIGHT, AGE,** AND **HEART RATE.** THE INVESTIGATOR IS INTERESTED IN THE RELATIONSHIPS BETWEEN THESE VARIABLES. HE DECIDES TO CALCULATE BASIC STATISTICS FOR ALL VARIABLES AND PEARSON CORRELATIONS BETWEEN ALL PAIRS OF VARIABLES. HE WISHES TO COMPUTE THE REGRESSION EQUATION BETWEEN HEART RATE AND WEIGHT, A PLOT BETWEEN THESE 2 VARIABLES, A FOURTH DEGREE POLYNOMIAL REGRESSION EQUATION BETWEEN HEART RATE AND HEIGHT, AND THE BEST REGRESSION EQUATION RELATING THE DEPENDENT VARIABLE, WEIGHT, WITH THE INDEPENDENT VARIABLES, HEIGHT AND AGE. A PLOT BETWEEN THESE 3 VARIABLES IS ALSO DESIRED. THE RAW DATA ARE SHOWN IN TABLE 2.

TABLE 2

SUBJECT	HEIGHT(IN)	WEIGHT(LBS)	AGE	HEART RATE
1	68	150	22	70
2	69	160	23	78
3	70	169	25	79
4	71	167	27	82
5	72	170	29	83
6	73	175	24	87
7	74	180	23	89
8	75	180	21	72

A LIST OF THE ASSEMBLED CARD DECK IS SHOWN IN FIGURE 16. FIGURES 17 THROUGH 29 SHOW THE COMPUTER OUTPUT. THIS EXAMPLE DEMONSTRATES HOW **$CMT** AND **$TITLE** CARDS MAY BE USED FOR LABELING.

EXAMPLE 3

24 TEST ANIMALS WERE RANDOMLY ASSIGNED IN GROUPS OF 8 TO 3 TREATMENTS, DIETS. THE ACTIVITY OF EACH ANIMAL WAS MEASURED. THE HYPOTHESIS OF NO DIFFERENCES BETWEEN DIETS COULD BE TESTED USING EITHER A ONE-WAY ANALYSIS OF VARIANCE OR THE KRUSKAL-WALLIS ROUTINE.

A PARAMETRIC OR A NON-PARAMETRIC MULTIPLE COMPARISON PROCEDURE COULD BE USED TO DECIDE WHICH TREATMENTS DIFFER. (AN INVESTIGATOR CONFRONTED WITH A SIMILAR PROBLEM SHOULD CHOOSE EITHER THE PARAMETRIC OR NON-PARAMETRIC APPROACH TO ANALYZING THE DATA. BOTH METHODS ARE DEMONSTRATED IN THIS EXAMPLE ONLY TO SHOW THE DIFFERENT OUTPUTS.) THE RAW DATA ARE PRESENTED IN TABLE 3.

TABLE 3

DIET 1	DIET 2	DIET 3
14	12	19
15	16	17
14	12	19
17	13	17
16	17	18
18	14	19
19	13	16
15	16	17

A LIST OF THE ASSEMBLED CARD DECK IS SHOWN IN FIGURE 30. FIGURES 31 THROUGH 37 SHOW THE OUTPUT FOR THIS EXAMPLE.

EXAMPLE 4

AN INVESTIGATOR PERFORMED TWO EXPERIMENTS. IN THE FIRST STUDY, 7 TEST ANIMALS WERE FED A CONTROL DIET AND THEIR ACTIVITIES WERE MEASURED. THE SAME 7 ANIMALS WERE THEN FED AN ENRICHED DIET AND THEIR ACTIVITIES WERE AGAIN MEASURED. IN THE SECOND STUDY 20 TEST ANIMALS WERE RANDOMLY ASSIGNED IN GROUPS OF 10 TO EITHER THE CONTROL DIET OR ENRICHED DIET. FOR REASONS UNRELATED TO EITHER TREATMENT, ONE ANIMAL IN THE CONTROL GROUP AND TWO ANIMALS IN THE TEST GROUP WERE NOT MEASURED. (THE INVESTIGATOR†S CHILD OPENED 3 CAGES AND THE ANIMALS ESCAPED. THEY WERE LATER RETRIEVED FROM SOME GARBAGE CANS WHERE THEY HAD BEEN FEASTING. NATURALLY, THEY HAD TO BE DISQUALTIFIED.) THE INVESTIGATOR WAS INTERESTED IN THE EFFECTS OF THE ENRICHED DIETS, IF ANY, ON ACTIVITY. HE MIGHT HAVE USED EITHER A PAIRED T-TEST OR WILCOXON†S SIGNED RANKS TEST TO EVALUATE THE DATA FROM THE FIRST STUDY, AND HE MIGHT HAVE USED EITHER A NON-PAIRED T-TEST OR WILCOXON†S RANK SUM TEST TO ANALYSE THE DATA FROM THE SECOND STUDY. THE OUTPUT INCLUDES ALL 4 TESTS. THE RAW DATA ARE PRESENTED IN TABLE 4.

TABLE 4

STUDY 1 CONTROL DIET	STUDY 1 ENRICHED DIET	STUDY 2 CONTROL DIET	STUDY 2 ENRICHED DIET
15	13	13	15
17	19	15	19
10	15	14	13
19	17	10	20
12	16	12	17
13	15	17	16
16	18	17	15
		18	18
		15	

A LISTING OF THE ASSEMBLED CARD DECK IS GIVEN IN FIGURE 38. FIGURES 39 THROUGH 47 SHOW THE OUTPUT FOR THIS EXAMPLE.

EXAMPLE 5

EXAMPLE 5 SHOWS SOME USES OF $TRANS AND $ADD. FIGURE 48 SHOWS THE ASSEMBLED CARD DECK. FIGURES 49 THROUGH 57 SHOW THE OUTPUT.

ALL SYSTEM CARDS IN THIS AND FOLLOWING SECTIONS ARE FOR THE **CDC7600 BKY** OPERATING SYSTEM. IN THE FOLLOWING EXAMPLES, THE **GRASS** CONTROL CARDS, (THOSE BEGINNING WITH A DOLLAR SIGN), INDICATE THAT THERE ARE 10 INPUT VARIABLES WITH 15 OBSERVATIONS, (CASES). THE INPUT DATA IS NOT TO BE PRINTED AND BASIC STATISTICS ARE TO BE CALCULATED FOR ALL VARIABLES.

THE FIRST CARD OF EVERY DECK IS THE **JOB CARD**. THE GENERAL ABBREVIATED FORMAT OF THE **JOB** CARD IS

JOBNAME.ACCNUM,YOUR NAME

WHERE

```
    JOBNAME - 1-7 ALPHANUMERIC CHARACTERS BEGINNING WITH A LETTER
              IN COLUMN 1
     ACCNUM - YOUR 6 DIGIT ACCOUNT NUMBER
  YOUR NAME - ANYTHING IN THE REMAINING COLUMNS.
```

LAIR USERS SHOULD CONSULT **MEMORANDUM FOR ALL UT200 USERS** DATED **AUGUST 16, 1974** FOR SPECIFIC CHARACTERS TO BE USED FOR **-JOBNAME-**, **-ACCNUM-**, AND **-YOUR NAME-**.

EXAMPLES

1. INPUT DATA ON CARDS

```
JOB CARD
FETCHPS(GRASZ,GRAZZ,GRASS)
GRAZZ.
(7/8/9 MULTIPUNCH CARD)
$TITLE EXAMPLE WITH INPUT DATA ON CARDS
$DATA  10,15,0
$FORMAT(10F5.0)
    DATA CARDS
$BSTAT ALL
$END
(6/7/8/9 MULTIPUNCH CARD)
```

2. INPUT DATA IN FORMATTED OR BINARY FORM ON MAGNETIC TAPE.

```
JOB CARD
FETCHPS(GRASZ,GRAZZ,GRASS)
STAGE,MYTAPE,TAPENO.
GRAZZ.
(7/8/9 MULTIPUNCH CARD)
$TITLE EXAMPLE WITH INPUT DATA ON MYTAPE
$DATA  10,15,0
$FILE  MYTAPE          OR        $FILEB MYTAPE
$FORMAT(10F5.0)                  NO FORMAT STATEMENT
                                 DATA ON MYTAPE MUST BE IN BINARY FORM
$BSTAT ALL
$END
(6/7/8/9 MULTIPUNCH CARD)
```

3. INPUT DATA IN FORMATTED OR BINARY FORM ON A PERMANENT FILE DISK.

```
JOB CARD
FETCHPS(GRASZ,GRAZZ,GRASS)
ATTACH(MYDATA,MYFILE)
GRAZZ.
(7/8/9 MULTIPUNCH CARD)
$TITLE EXAMPLE WITH INPUT DATA ON DISK
$DATA   10,15,0
$FILE   MYDATA            OR        $FILEB MYDATA
$FORMAT(10F5.0)                     NO FORMAT STATEMENT
                                    DATA ON MYDATA MUST BE IN BINARY FORM
$BSTAT ALL
$END
(6/7/8/9 MULTIPUNCH CARD)
```

4. INPUT DATA IS ON CARDS IN FORMATTED FORM. TWO NEW VARIABLES ARE CREATED USING $TRANS CARDS. THESE NEW VARIABLES ARE WRITTEN IN FORMATTED OR BINARY FORM TO A MAGNETIC TAPE.

```
JOB CARD
FETCHPS(GRASZ,GRAZZ,GRASS)
GRAZZ.
STAGE,NEWVAR,TAPENO,W.
(7/8/9 MULTIPUNCH CARD)
$TITLE WRITING NEW VARIABLES TO NEWVAR.
$DATA   10,15,0
$TRANS X(11)=2* X(1)
$TRANS X(12)=20-X(11)
$FORMAT(10F5.0)
     DATA CARDS
$BSTAT ALL
$WRITE 11,12
$FILE   NEWVAR            OR        $FILEB NEWVAR
$FORMAT(2F5.0)                      NO FORMAT CARD
$END
(6/7/8/9 MULTIPUNCH CARD)
```

IF NONE OF THESE EXAMPLES SUITS YOUR PROBLEM, OR MACHINE, OR IF YOU ARE UNSURE ABOUT THE APPROPRIATE SYSTEM CARDS TO USE, SPEAK TO A LOCAL CONSULTANT.

TITLE

```
1      8
$TITLE A TITLE
```

PARAMETER FIELD (COLUMNS 8-79)
ANY ALPHANUMERIC TITLE IN COLUMNS 8-79

OUTPUT
THE TITLE IS PRINTED ON THE TOP OF ALL PAGES OF OUTPUT UNTIL ANOTHER $TITLE CARD IS USED.

USAGE
1. THE TITLE MUST BE CONTAINED ON ONE $TITLE CARD.
2. COLUMN 80 MUST NOT BE PUNCHED ON THIS CARD.

EXAMPLE

```
1      8
$TITLE ANALYSES OF MY DATA
```

DESCRIBING INPUT DATA

```
1       8
$DATA   I1,I2,I3
```

PARAMETER FIELD (COLUMNS 8-79)

- I1 NUMBER OF VARIABLES
- I2 NUMBER OF CASES
- I3 0-DO NOT PRINT INPUT DATA
  1-PRINT INPUT DATA

OUTPUT

1. INPUT DATA IS PRINTED IF I3=1.

USAGE

1. MAXIMUM OF 50 INPUT VARIABLES (I1≤50)
2. NO LIMIT ON NUMBER OF OBSERVATIONS

EXAMPLES

CARDS		COMMENTS
1	8	
$DATA	6,30,1	6 VARIABLES, 30 OBSERVATIONS PRINT INPUT DATA
$DATA	7,10,0	7 VARIABLES, 10 OBSERVATIONS DO NOT PRINT INPUT DATA

WRITING SPECIFIED VARIABLES

```
1      8
$WRITE I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..  VARIABLES TO BE WRITTEN
    ..
    I50
```

RESULT

1. IF $WRITE IS IMMEDIATELY FOLLOWED BY A $FORMAT CARD, ALL VALUES (CASES) OF THE VARIABLES SPECIFIED IN THE PARAMETER FIELD WILL BE WRITTEN ACCORDING TO THE FORMAT SPECIFIED ON THE $FORMAT CARD TO THE LOCAL FILE BCDOUT.
2. IF $WRITE IS FOLLOWED BY A $FILE CARD AND A $FORMAT CARD, ALL CASES OF THE VARIABLES SPECIFIED WILL BE WRITTEN TO THE FILE NAMED ON THE $FILE CARD IN THE FORMAT SPECIFIED ON THE $FORMAT CARD.
3. IF $WRITE IS FOLLOWED BY A $FILEB CARD, ALL CASES OF THE THE VARIABLES SPECIFIED WILL BE WRITTEN IN BINARY, (NON-FORMATTED), FORM TO THE FILE NAMED ON THE $FILEB CARD. NO $FORMAT CARD IS USED WITH $FILEB. IF ONE IS USED IT WILL BE IGNORED.

USAGE

1. A MAXIMUM OF 50 VARIABLES MAY BE WRITTEN.
2. NO LIMIT ON NUMBER OF CASES WRITTEN.

EXAMPLES

CARDS	COMMENTS
1 8	
$WRITE 1,3,5 $FORMAT(3F5.0)	VARIABLES 1,3, AND 5 WILL BE WRITTEN TO BCDOUT IN 3 FIELDS OF F5.0.
$WRITE 1,3,5 $FILE MYDATA $FORMAT(3F5.0)	VARIABLES 1,3 AND 5 WILL BE WRITTEN TO MYDATA IN 3F5.0 FORMAT.
$WRITE 1,3,5 $FILEB BINDATA	VARIABLES 1,3 AND 5 WILL BE WRITTEN TO BINDATA IN BINARY (UNFORMATTED) FORM.

CODED FILES

```
1      8
$FILE  LFNAME
```

PARAMETER FIELD (COLUMNS 8-14)

1. THE LOCAL FILE NAME OF THE CODED, FORMATTED, FILE THAT CONTAINS THE INPUT DATA, OR
2. THE LOCAL FILE NAME OF THE FILE THE DATA IS TO BE WRITTEN TO IN FORMATTED FORM.

USAGE

1. WHEN READING FORMATTED DATA FROM A SOURCE OTHER THAN CARDS, THE $FILE CARD MUST COME BETWEEN THE $DATA AND $FORMAT CARDS.
2. WHEN WRITING FORMATTED DATA TO AN ALTERNATE OUTPUT DEVICE, THE $FILE CARD MUST COME BETWEEN THE $WRITE AND $FORMAT CARDS. IF NO LOCAL FILE NAME IS PUNCHED IN THE PARAMETER FIELD OF THE $FILE CARD, DATA SPECIFIED WILL BE WRITTEN TO BCDOUT.

RESULT

1. INPUT DATA WILL BE READ FROM THE FILE NAMED, OR
2. SPECIFIED DATA WILL BE WRITTEN TO THE FILE NAMED.

EXAMPLES

1. FOR READING DATA FROM AN ALTERNATE INPUT SOURCE SEE APPENDIX A, EXAMPLES 2 AND 3.
2. FOR WRITING DATA TO AN ALTERNATE OUTPUT DEVICE SEE APPENDIX A, EXAMPLE 4.

BINARY FILES

```
1      8
$FILEB LFNAME
```

PARAMETER FIELD (COLUMNS 8-14)
1. THE LOCAL FILE NAME OF THE BINARY, (UNFORMATTED), FILE THAT CONTAINS THE INPUT DATA, OR
2. THE LOCAL FILE NAME OF THE BINARY FILE THAT SPECIFIED VARIABLES ARE TO BE WRITTEN TO IN BINARY FORM.

USAGE
1. WHEN READING DATA FROM BINARY FILE, THE $FILEB CARD FOLLOWS THE $DATA CARD.
2. WHEN WRITING A BINARY FILE, THE $FILEB CARD FOLLOWS THE $WRITE CARD.
3. NO $FORMAT CARD IS USED WITH THE $FILEB CARD.

RESULT
DATA MAY BE READ FROM OR WRITTEN TO A BINARY FILE.

EXAMPLE
1. TO READ DATA FROM A BINARY FILE, SEE APPENDIX A, EXAMPLES 2 AND 3.
2. TO WRITE A BINARY FILE, SEE APPENDIX A, EXAMPLE 4.

FORMAT STATEMENT

```
1     8
$FORMAT(FORTRAN FORMAT STATEMENT)
```

PARAMETER FIELD (COLUMNS 8-79)

ANY FORTRAN FORMAT STATEMENT USING F, E, AND/OR X SPECIFICATIONS

RESULT

DESCRIBES THE FIELDS OF INPUT DATA.

USAGE

1. A FORMAT STATEMENT MAY BE CONTINUED ON FOLLOWING CARDS BUT THE TOTAL LENGTH CANNOT EXCEED 250 CHARACTERS.
2 THE $FORMAT CARD MUST FOLLOW THE $DATA CARD IF INPUT DATA IS ON CARDS IN FORMATTED FORM.
3 THE $FORMAT CARD FOLLOWS THE $FILE CARD IF INPUT DATA IS ON AN ALTERNATE INPUT SOURCE IN FORMATTED FORM OR IF VARIABLES ARE TO BE WRITTEN IN FORMATTED FORM TO AN ALTERNATE OUTPUT DEVICE.
4. $FORMAT IS NOT USED WITH THE $FILEB CARD.

EXAMPLES

CARDS

```
1     8
$FORMAT(2F3.2,1X,F4.2)
$FORMAT(F4.1,F3.0,F4.2,F2.0)
```

FORMFREE INPUT DATA

```
1     8
$FMFREE
```

PARMETER FIELD (COLUMNS 8-79)
    NO ENTRY

RESULT
    DATA PUNCHED IN A FORMAT-FREE MANNER IS READ.

USAGE
1. DATA VALUES ARE PUNCHED ON CARDS IMMEDIATELY FOLLOWING $FMFREE CARD.
2. EACH DATA VALUE MAY BE PUNCHED IN EITHER F-TYPE OR E-TYPE FORMAT WITH AS MANY DIGITS AS DESIRED, BUT A DECIMAL POINT MUST BE PUNCHED.
3. DATA VALUES MUST BE SEPARATED BY BLANK SPACES OR COMMAS.
4. AS MANY VALUES PER CARD AS DESIRED MAY BE PUNCHED, BUT A DATA VALUE MUST BE PUNCHED ENTIRELY ON ONE CARD.
5. DATA VALUES FOR EACH OBSERVATION MUST BEGIN ON A NEW CARD.

COMMENT CARD

```
1      8
$CMT   ANY COMMENT
```

PARAMETER FIELD (COLUMNS 8-79)
ANY COMMENT

OUTPUT
THE COMMENT IN COLUMNS 8-79 WILL BE PRINTED ON THE NEXT PAGE OF OUTPUT. IF TWO OR MORE $CMT CARDS ARE IN SEQUENCE, ALL COMMENTS WILL BE PRINTED ON THE NEXT PAGE USING TRIPLE SPACING.

USAGE
COLUMN 80 MUST NOT BE PUNCHED ON THE $CMT CARD.

EXAMPLE

```
COLUMN
1      8
$CMT   VARIABLE 3 IS HEART RATE.
```

REWIND

```
1      8
$RWD   LFNAME
```

PARAMETER FIELD (COLUMNS 8-14)
THE LOCAL FILE NAME OF THE FILE TO BE REWOUND

RESULT
THE FILE SPECIFIED WILL BE REWOUND.

USAGE
1. ONLY ONE FILE MAY BE REWOUND WITH ANY ONE $RWD CARD.
2. THE LOCAL FILE NAMED MUST REFER TO THE FILE IN CURRENT USE.
3. DO NOT REWIND TAPE3 OR TAPE4.

EXAMPLES

CARDS		COMMENTS
1	8	
$RWD	MYDATA	MYDATA WILL BE REWOUND.
$RWD	MYTAPE	MYTAPE WILL BE REWOUND.

HISTOGRAM SCALING

```
1      8
$SCALE Y1,Y2
```

PARAMETER FIELD (COLUMNS 8-79)

Y1 - MINIMUM
Y2 - MAXIMUM

RESULT

SETS THE MINIMUM AND MAXIMUM FOR THE BASE AXIS, AND THEREBY DETERMINES THE SCALE FACTOR FOR THE HISTOGRAM ROUTINE.

USAGE

1. IF THE $SCALE CARD IS USED, IT MUST PRECEED THE $HSTGM CARD.
2. ABSENCE OF THE $SCALE CARD WILL CAUSE AUTO-SCALING.

EXAMPLES

CARDS	COMMENTS
1 8	
$SCALE 20, 160	SCALE FROM 20 TO 160.
$SCALE 10.3, 100.3	SCALE FROM 10.3 TO 100.3.

PLOT SCALING

```
1      8
$SCALE Y1,Y2,X1,X2
```

PARAMETER FIELD (COLUMNS 8-79)

Y1 - MINIMUM FOR THE DEPENDENT VARIABLE (Y-AXIS)
Y2 - MAXIMUM FOR THE DEPENDENT VARIABLE (Y-AXIS)
X1 - MINIMUM FOR THE INDEPENDENT VARIABLE (X-AXIS)
X2 - MAXIMUM FOR THE INDEPENDENT VARIABLE (X-AXIS)

RESULT

SETS THE MINIMUM AND MAXIMUM FOR THE X AND/OR Y AXIS, AND THEREBY DETERMINES THE SCALE FACTORS FOR THE PLOT ROUTINES.

USAGE

1. WHEN USED, THE $SCALE CARD MUST PRECEED EITHER THE $PLOT, $PLOTN, OR $PLOTI CARD.
2. ABSENCE OF THE $SCALE CARD WILL CAUSE AUTO-SCALING.
3. IF ONLY ONE AXIS IS TO BE MANUALLY SCALED, THEN THE MINIMUM AND MAXIMUM VALUES OF THE AUTO SCALED AXIS MUST BOTH BE SET TO ZERO.

EXAMPLES

CARDS	COMMENTS
1 8	
$SCALE 0,0,10,300	AUTO SCALING ON THE Y-AXIS, SCALING FROM 10 TO 300 ON THE X-AXIS
$SCALE 30,100,4.3,6.9	SCALING FROM 30 TO 100 ON THE Y-AXIS, SCALING FROM 4.3 TO 6.9 ON THE X-AXIS

COMBINING VARIABLES - $ADD

```
1      8
$ADD   I1,I2,I3
```

PARAMETER FIELD (COLUMNS 8-79)
- I1 - VARIABLE WHICH IS TO BE ON THE TOP
- I2 - VARIABLE WHICH IS TO BE ON THE BOTTOM
- I3 - CREATED NEW VARIABLE

RESULT

TWO VARIABLES, I1 AND I2, ARE VERTICALLY JOINED IN A NEW VARIABLE, I3.

USAGE

1. THE BOTTOM VARIABLE, I2, MAY CONTAIN NO MORE THAN 200 OBSERVATIONS, (CASES), IGNORING BLANKS.
2. OLD VARIABLES CAN AND WILL BE REWRITTEN WITH INCORRECT USE OF I3.

EXAMPLE

SUPPOSE THERE ARE 2 VARIABLES WITH THE VALUES

```
1 5
2 6
3 7
4 8
```

$ADD 1,2,3 WOULD RESULT IN THE CREATION OF A 3RD VARIABLE WITH VALUES

```
1
2
3
4
5
6
7
8
```

END

```
1      8
$END
```

PARAMETER FIELD
NO ENTRY

RESULT
$END INDICATES THAT THE CURRENT GRASS RUN IS FINISHED.

USAGE
THE $END CARD MUST BE THE LAST GRASS CONTROL CARD.

BASIC STATISTICS

```
1      8
$BSTAT I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1
I2
..
..   INPUT VARIABLES
..
I50
```

OUTPUT

1. VARIABLES
2. OBSERVATIONS
3. MINIMUM
4. MAXIMUM
5. RANGE
6. MEDIAN
7. MEAN
8. STANDARD DEVIATION
9. ST. ERROR OF THE MEAN

USAGE

MAXIMUM OF 50 VARIABLES
NO LIMIT ON NUMBER OF OBSERVATIONS
THE MEDIAN WILL NOT BE CALCULATED IF OBSERVATIONS>500.

EXAMPLE

CARDS	COMMENTS
1 8	
$BSTAT 4,5,9,10	BASIC STATISTICS FOR VARIABLES 4,5,9 AND 10
$BSTAT ALL	BASIC STATISTICS FOR ALL VARIABLES

FOR EACH VARIABLE REQUESTED THE BSTAT ROUTINE COMPUTES A VARIETY OF DESCRIPTIVE STATISTICS. THESE INCLUDE

1. MINIMUM - THE OBSERVATION WITH THE SMALLEST VALUE

2. MAXIMUM - THE OBSERVATION WITH THE LARGEST VALUE

3. RANGE - THE MAXIMUM VALUE MINUS THE MINIMUM VALUE

4. MEDIAN - THE VALUE FOR WHICH 50¬ OF THE VALUES ARE GREATER AND 50¬ ARE LESS. FOR AN ODD NUMBER OF OBSERVATIONS, THE MEDIAN IS THE MIDDLE VALUE IN A SEQUENTIALLY ORDERED SET. FOR AN EVEN NUMBER OF OBSERVATIONS, THE MEDIAN IS THE AVERAGE OF THE TWO MIDDLE VALUES.

5. MEAN - THE AVERAGE OF THE OBSERVATIONS

6. STANDARD DEVIATION - THE STANDARD DEVIATION IS DEFINED AS THE SQUARE ROOT OF THE SUM OF THE SQUARED DEVIATION OF EACH OBSERVATION FROM THE MEAN DIVIDED BY THE NUMBER OF OBSERVATIONS LESS 1.

SD= SQRT[SUM((X(I)-MEANX)↑2)/(N-1)]

WHERE- X(I) IS AN OBSERVATION, MEANX IS THE OVERALL MEAN, ↑ MEANS RAISED TO THE FOLLOWING POWER, N IS THE NUMBER OF OBSERVATIONS AND SUM INDICATES THE SUM OVER ALL OBSERVATIONS. THE STANDARD DEVIATION IS COMPUTED BY THE FORMULA

$$V = \frac{SUM(X(I)\uparrow 2)-[(SUM\ X(I))\uparrow 2]/N}{N-1}$$

SD = SQRT(V)

V IS THE VARIANCE.

7. STANDARD ERROR OF THE MEAN - STD. ERROR = SD/SQRT(N) - WHERE SD IS THE STANDARD DEVIATION AND N IS THE NUMBER OF OBSERVATIONS.

NON-PAIRED T-TEST

```
1      8
$NTST  I1,I2,...,I50
```

PARAMETER FIELD COLUMNS(8-79)

```
    I1
    I2
    ..
    ..   INPUT VARIABLES
    ..
    I50
```

OUTPUT

1. MEANS, STANDARD DEVIATIONS, AND NUMBER OF OBSERVATIONS, (CASES), FOR ALL INPUT VARIABLES ARE PRINTED.
2. T-VALUES, DEGREES OF FREEDOM, AND LEVELS OF SIGNIFICANCE OF T-VALUES FOR ALL PAIRS OF INPUT VARIABLES ARE PRINTED.

USAGE

A MAXIMUM OF 50 VARIABLES MAY BE USED.
THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

CARDS	COMMENTS
1 8	
$NTST 3,5	A NON-PAIRED T-TEST BETWEEN VARIABLES 3 AND 5 WILL BE PERFORMED.
$NTST ALL	NON-PAIRED T-TESTS BETWEEN ALL PAIRS OF INPUT VARIABLES WILL BE PERFORMED.

IN GENERAL, THE NON-PAIRED T-TEST IS USED TO TEST HYPOTHESES CONCERNING THE MEANS OF 2 POPULATIONS. THE USUAL HYPOTHESIS IS THAT THE MEANS OF THE 2 POULATIONS ARE EQUAL. A RANDOM SAMPLE IS CHOSEN FROM EACH OF THE 2 POPULATIONS. THE SAMPLE MEANS ARE USED TO MAKE INFERENCES ABOUT THE POPULATION MEANS. IT IS APPROPRIATE TO USE THE NON-PAIRED T-TEST TO HELP MAKE THESE INFERENCES WHEN THE DESIGN OF THE EXPERIMENT INCLUDES SAMPLING FROM ONLY 2 POPULATIONS. IF INFERENCES ARE TO BE MADE ABOUT THE MEANS OF MORE THAN 2 POPULATIONS, THE ANALYSIS OF VARIANCE OR THE KRUSKAL-WALLIS TEST SHOULD BE USED.

A TEST STATISTIC, T, IS CALCULATED WHICH MAY BE USED TO TEST THE HYPOTHESIS THAT THE TRUE MEAN OF POPULATION 1, MU1, EQUALS THE TRUE MEAN OF POPULATION 2, MU2, THE TEST STATISTIC IS CALCULATED AS

T= (MX1-MX2)/SQRT(V/N1 + V/N2)

WHERE MX1,MX2 ARE THE MEANS OF SAMPLES 1 AND 2,
N1,N2 ARE THE NO. OF OBSERVATIONS IN SAMPLES 1 AND 2,
AND V IS THE POOLED VARIANCE OF THE 2 SAMPLES.
THE FORMULA FOR -V- IS

V= [(N1-1)V1 + (N2-1)V2]/N1+N2-2

WHERE V1 AND V2 ARE THE VARIANCES OF SAMPLES 1 AND 2. T HAS DEGREES OF FREEDOM (D.F.) = N1+N2-2. THE LEVEL OF SIGNIFICANCE, ALPHA, PRINTED IN THE OUTPUT IS EQUAL TO TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T. IN OTHER WORDS, ALPHA IS THE SMALLEST SIGNIFICANCE LEVEL AT WHICH THE HYPOTHESIS MU1=MU2 COULD BE REJECTED IN FAVOR OF THE TWO-SIDED ALTERNATIVE HYPOTHESIS, MU1≠MU2. FOR THE SMALLEST SIGNIFICANCE LEVEL FOR WHICH THE HYPOTHESIS MU1=MU2 COULD BE REJECTED IN FAVOR OF EITHER OF THE ONE-SIDED ALTERNATE HYPOTHESES,MU1<MU2 OR MU1>MU2, USE ALPHA/2.

REFERENCE

STEELE, R.G.D. AND J.H. TORRIE (1960) **PRINCIPLES AND PROCEDURES OF STATISTICS** MCGRAW-HILL, NEW YORK. PP 73-78.

PAIRED T-TEST

```
1      8
$PTST  I1,I2,....I25
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..   INPUT VARIABLES
    ..
    I25
```

OUTPUT

1. MEANS, STANDARD DEVIATIONS, AND NUMBERS OF OBSERVATIONS ARE PRINTED FOR ALL INPUT VARIABLES.
2. T-VALUES, DEGREES OF FREEDOM, AND LEVELS OF SIGNIFICANCE OF T-VALUES ARE PRINTED FOR ALL PAIRS OF INPUT VARIABLES.

USAGE

A MAXIMUM OF 25 VARIABLES MAY BE USED.
THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS, (CASES).

EXAMPLES

CARDS		COMMENTS
1	8	
$PTST	4,6	A PAIRED T-TEST BETWEEN VARIABLES 4 AND 6 WILL BE PERFORMED.
$PTST	ALL	PAIRED T-TESTS BETWEEN ALL PAIRS OF INPUT VARIABLES WILL BE PERFORMED.

PAIRED T-TESTS ARE USED TO TEST THE HYPOTHESIS THAT THE MEAN OF THE DIFFERENCES, MUD, BETWEEN TWO POPULATIONS IS 0. THE SAMPLES MUST BE PAIRED. SOME EXAMPLES OF PAIRED SAMPLES ARE BEFORE AND AFTER TREATMENTS ON THE SAME SUBJECT OR SAMPLES IN WHICH SUBJECTS ARE PAIRED ON SOME CRITERIA OTHER THAN THAT BEING MEASURED. THE DECISICION OF WHETHER OR NOT TO USE PAIRING MUST BE MADE BEFORE THE EXPERIMENT IS CONDUCTED. PAIRED T-TESTS SHOULD NOT BE USED UNLESS THE SAMPLES WERE IN FACT PAIRED.

THE TEST STATISTIC, T, IS DEFINED AS

T= MD/SQRT(VD/N)

WHERE MD IS THE AVERAGE OF THE DIFFERENCES BETWEEN SAMPLES 1 AND 2,
VD IS THE VARIANCE OF THE DIFFERENCES,
AND N IS THE NUMBER OF PAIRS OR DIFFERENCES.
T HAS DEGREES OF FREEDOM (D.F.) EQUAL TO N-1.

TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T VALUE IS PRINTED AS THE LEVEL OF SIGNIFICANCE, ALPHA. THIS IS THE SMALLEST SIGNIFICANCE LEVEL AT WHICH THE HYPOTHESIS, MUD=0, MAY BE REJECTED IN FAVOR OF THE TWO-SIDED ALTERNATE HYPOTHESIS MUD≠0. FOR THE SMALLEST SIGNIFICANCE LEVEL AT WHICH THE HYPOTHESIS MUD=0 COULD BE REJECTED IN FAVOR OF EITHER OF THE ONE-SIDED ALTERNATE HYPOTHESES, MUD<0 OR MUD>0, USE ALPHA/2.

REFERENCE

STEELE, R.G.D. AND J.H. TORRIE (1960) **PRINCIPLES AND PROCEDURES OF STATISTICS.** MCGRAW-HILL, NEW YORK. PP 78-80.

WILCOXON*S RANK SUM TEST

```
1     8
$RNKSM I1,I2
```

PARAMETER FIELD (COLUMNS 8-79)

I1 - THE FIRST VARIABLE
I2 - THE SECOND VARIABLE

OUTPUT

1. THE RELATIVE RANKS OF THE SECOND INPUT VARIABLE, I2, AND THE SUM OF THOSE RANKS, THE WILCOXON STATISTIC, W, ARE PRINTED.
2. WHEN THE TOTAL NUMBER OF OBSERVATIONS IN I1 AND I2 IS LESS THAN 16, UPPER TAIL PROBABILITIES FOR -W- AND FOR *N(M+N+1)-W* ARE PRINTED WHERE M= NO. OF OBSERVATIONS OF THE FIRST VARIABLE, I1, AND N= NO. OF OBSERVATIONS OF THE SECOND VARIABLE, I2.
3. THE LARGE SAMPLE APPROXIMATION TO -W- AND THE LEVEL OF SIGNIFICANCE OF THE LARGE SAMPLE APPROXIMATION ARE PRINTED.
4. THE RELATIVE RANKS OF THE SECOND INPUT VARIABLE, I2, FOLLOWING A RERANKING FOR THE SIEGEL-TUKEY TEST ARE OUTPUT. THE SUM OF THESE RANKS IS ALSO OUTPUT.

USAGE

1. TO AVOID DIFFICULTIES IN THE INTERPRETATION OF THE OUTPUT, THE VARIABLE WITH THE LESSER NUMBER OF OBSERVATIONS SHOULD BE INPUT AS THE SECOND VARIABLE, I2.
2. THE TOTAL NUMBER OF OBSERVATIONS MAY NOT EXCEED 1000, AND EACH VARIABLE MAY CONTAIN NO MORE THAN 500 OBSERVATIONS.

EXAMPLE

CARD	COMMENTS
1 8	
$RNKSM 9,7	WILCOXON*S RANK SUM TEST WILL BE PERFORMED FOR VARIABLES 9 AND 7.

I. THE WILCOXON RANK SUM STATISTIC IS USED TO TEST HYPOTHESES ABOUT THE DIFFERENCE BETWEEN THE MEDIANS OF TWO POPULATIONS. IN PARTICULAR, THE NULL HYPOTHESIS IS MED(P1)-MED(P2)=DELTA=0, WHERE MED(P1) AND MED(P2) ARE THE MEDIANS OF POPULATIONS 1 AND 2.

SUPPOSE WE HAVE OBTAINED 2 SAMPLES OF K=M+N OBSERVATIONS, X1,...., XM AND Y1,...., YN. FIRST THE K OBSERVATIONS ARE ORDERED FROM LEAST TO GREATEST. THE ORDERED OBSERVATIONS ARE ASSIGNED RANKS 1 TO K. TIED OBSERVATIONS ARE ASSIGNED THE AVERAGE RANK OF THE TIED OBSERVATIONS. THE SUM OF THE N RANKS OF THE Y VARIABLE IS THE RANK SUM STATISTIC, W. THE SMALLEST SIGNIFICANCE LEVEL, ALPHA1, AT WHICH THE HYPOTHESIS, MED(PY)-MED(PX)=DELTA=0, COULD BE REJECTED IN FAVOR OF THE ALTERNATE HYPOTHESIS, DELTA>0 IS PRINTED AS THE UPPER TAIL PROBABILITY OF W. THE SMALLEST SIGNIFICANCE, ALPHA2, AT WHICH THE HYPOTHESIS, DELTA=0 COULD BE REJECTED IN FAVOR OF THE ALTERNATE HYPOTHESIS DELTA<0 IS PRINTED AS THE UPPER TAIL PROBABILITY OF N(M+N+1)-W. FOR THE LEVEL OF SIGNIFICANCE AGAINST THE TWO-SIDED ALTERNATE HYPOTHESIS, DELTA≠0, CALCULATE ALPHA=ALPHA1+(1-ALPHA2) IF ALPHA2>ALPHA1 OR ALPHA=(1-ALPHA1)+ALPHA2 IF ALPHA1>ALPHA2.. IF W IS NOT A WHOLE NUMBER, IT IS ROUNDED TO THE NEXT HIGHEST WHOLE NUMBER FOR CALCULATION OF ALPHA1, AND N(N+M+1)-W IS ROUNDED TO THE NEXT LOWEST WHOLE NUMBER FOR CALCULATION OF ALPHA2.

FOR EASE OF INTERPRETATION ALWAYS USE N≤M. THIS CAN BE ACCOMPLISHED BY DESIGNATING THE VARIABLE WITH THE SMALLER NUMBER OF OBSERVATIONS AS THE SECOND INPUT VARIABLE, I2.

ALL OUTPUT INCLUDES THE LARGE SAMPLE APPROXIMATION TO W, W*. FOR LARGE K, W* FOLLOWS A STANDARD NORMAL DISTRIBUTION. THE GENERAL FORMULA FOR W* IS

$$W^* = \frac{W-[(N+M+1)/2]}{SQRT(VAR(W))} .$$

IN GENERAL,

$$VAR(W)=(MN/12)[M+N+1 - SUM(T(J))(T(J)\uparrow 2-1)/(M+N)(M+N+1)]$$

WHERE T(J)= THE NUMBER OF OBSERVATIONS IN THE JTH GROUP OF TIED OBSERVATIONS AND SUM INDICATES SUMMING (T(J))(T(J)↑2-1) OVER ALL J GROUPS. FOR UNTIED OBSERVATIONS, T(J)=1. WHEN THERE ARE NO TIED OBSERVATIONS, VAR(W) REDUCES TO

$$VAR(W) = MN(M+N+1)/12$$

AND

$$W^* = \frac{W-[N(M+N+1)/2]}{SQRT(MN(M+N+1)/12)}$$

THE LEVEL OF SIGNIFICANCE OF W* IS THE PROBABILITY OF FINDING A RANDOM VALUE OF W* GREATER THAN THE OBTAINED VALUE OF W*.

THE WILCOXON RANK SUM TEST MAY BE USED INSTEAD OF THE UNPAIRED T-TEST TO DISCERN INFORMATION ABOUT THE EQUALITY OF THE CENTERS OF TWO POPULATIONS.

II. THE SIEGEL-TUKEY TEST MAY BE USED TO TEST HYPOTHESES ABOUT THE

DIFFERENCES IN THE SPREAC OR VARIANCES OF TWO POPULATIONS. FOR EXAMPLE, IT MIGHT BE USED TO TEST THE HYPOTHESIS THAT THE PRECISION OF TWO MEASURING INSTRUMENTS IS THE SAME. IN GENERAL, THE NULL HYPOTHESIS IS **VAR(PY)=VAR(PX)**. THE **K=M+N** OBSERVATIONS ARE CRDERED FROM LEAST TO GREATEST. RANKS ARE THEN ASSIGNED BY GIVING RANK 1 TO THE SMALLEST OBSERVATION, RANK 2 TC THE LARGEST, RANK 3 TO THE SECOND LARGEST, RANK 4 TO THE SECOND SALLEST AND SO ON. THE ORDERED **M+N** OBSERVATIONS THEN HAVE RANKS

1,4,5,8,9,....,(M+N),....,7,6,3,2.

TIED OBSERVATIONS ARE ASSIGNED THE AVERAGE RANK OF THE TIED OBSERVATIONS. AS IN THE RANK SUM TEST, THE SUM OF THE **N** RANKS OF THE **Y** VARIABLE IS THE TEST STATISTIC, **W**. A PRIMARY ASSUMPTION OF THE SIEGEL-TUKEY TEST IS THAT THE CENTERS OF THE TWO POPULATIONS ARE COINCIDENT. IF THERE IS REASON TO BELIEVE THAT THE CENTERS ARE NOT COINCIDENT, A TRANSFORMATION MAY BE PERFORMED ON THE OBSERVATIONS TO EQUALIZE THE MEDIANS. CAUTION SHOULD BE USED SINCE ADJUSTMENT OF THE DATA ALTERS THE NULL DISTRIBUTION OF **W**. IT IS PROBABLY BETTER TO USE THE **F-TEST** FOR COMPARISONS OF 2 VARIANCES WHEN THERE IS REASON TO BELIEVE THAT THE POPULATIONS DO NOT HAVE THE SAME CENTER.

REFERENCES

HOLLANDER, M. AND D.A. WOLFE (1973) **NON-PARAMETRIC STATISTICAL METHODS.** WILEY, NEW YORK. PP 67-75

CONOVER, W.J. (1971) **PRACTICAL NONPARAMETRIC STATISTICS.** WILEY, NEW YORK.

WILCOXON'S SIGNED RANKS TEST

```
1     8
$SNRNK I1,I2
```

PARAMETER FIELD (COLUMNS 8-79)

I1

INPUT VARIABLES TO BE COMPARED

I2

OUTPUT

1. THE RANKS OF THE ABSOLUTE VALUES OF THE DIFFERENCES BETWEEN THE PAIRED VARIABLES ARE PRINTED.
2. THE SUM OF THE RANKS CORRESPONDING TO POSITIVE DIFFERENCES, T+, THE SUM OF THE RANKS CORRESPONDING TO NEGATIVE DIFFERENCES, T-, AND THE MINIMUM OF THESE TWO SUMS, T, ARE OUTPUT. THE LEVEL OF SIGNIFICANCE OF T IS ALSO OUTPUT.
3. THE LARGE SAMPLE APPROXIMATION OF -T- AND THE LEVEL OF SIGNIFICANCE OF THE LARGE SAMPLE APPROXIMATION ARE PRINTED.

USAGE

1. A MAXIMUM OF 500 OBSERVATIONS, DIFFERENCES, MAY BE USED.
2. THE NUMBER OF OBSERVATIONS FOR THE TWO INPUT VARIABLES MUST BE EQUAL.

EXAMPLE

CARD	COMMENTS
1 8	
$SNRNK 3,5	A SIGNED RANKS TEST WILL BE PERFORMED FOR VARIABLES 3 AND 5.

WILCOXON'S SIGNED RANK TEST IS USED TO TEST HYPOTHESES ABOUT THE MEDIAN OF THE DIFFERENCES OF 2 PAIRED POPULATIONS.

CONSIDER A SET OF PAIRED OBSERVATIONS, **(X1,Y1), (X2,Y2),....,(XN, YN)**. FROM THIS SET N DIFFERENCES ARE CALCULATED?

```
D1 = X1-Y1
D2 = X2-Y2
 .    .  .
 .    .  .
 .    .  .
DN = XN-YN.
```

THE NULL HYPOTHESIS OF NO DIFFERENCES BETWEEN THE PAIRED OBSERVATIONS MAY BE FORMULATED AS **MED(D)=0**, WHERE **MED(D)** IS THE MEDIAN OF THE DIFFERENCES. ALTERNATIVE HYPOTHESES ARE **MED(D)≠0, MED(D)>0**, OR **MED(D)<0**.

TO TEST THE NULL HYPOTHESIS, THE ABSOLUTE VALUES OF THE DIFFERENCES ARE ORDERED FROM LEAST TO GREATEST. RANKS OF 1 TO N ARE THEN ASSIGNED TO THE ORDERED VALUES. THE SUM OF THE RANKS OF THE POSITIVE DIFFERENCES, T+, AND THE SUM OF THE RANKS OF THE NEGATIVE DIFFERENCES, T-, ARE COMPUTED. **TMIN**, THE MINIMUM OF **T+** AND **T-**, MAY BE USED AS THE TEST STATISTIC WHEN THE ALTERNATE HYPOTHESIS IS **MED(D)≠0**. THE PROBABILITY OF FINDING A RANDOM VALUE OF **T** LESS THAN OR EQUAL TO THE OBTAINED **TMIN** IS WRITTEN AS THE LEVEL OF SIGNIFICANCE. THIS IS EQUIVALENT TO THE PROBABILITY OF FINDING A RANDOM VALUE OF **T** GREATER THAN OR EQUAL TO **TMAX**, WHERE **TMAX** IS THE MAXIMUM OF **T+** AND **T-**. THUS TO OBTAIN THE SMALLEST SIGNIFICANCE LEVEL AT WHICH THE HYPOTHESIS **MED(D)=0** MAY BE REJECTED IN FAVOR OF THE ALTERNATE HYPOTHESIS **MED(D)≠0** USE TWICE THE VALUE OF THE SIGNIFICANCE LEVEL PRINTED. IF THE ALTERNATE HYPOTHESIS OF INTEREST IS **MED(D)>0**, FIND THE PROBABILITY OF FINDING A RANDOM VALUE OF **T** GREATER THAN OR EQUAL TO THE OBTAINED **T+**, IF THE ALTERNATE HYPOTHESIS IS **MED(D)<0**, FIND THE PROBABILITY OF FINDING A RANDOM VALUE OF **T** GREATER THAN OR EQUAL TO THE OBTAINED VALUE OF **T-**. THE ABOVE PROBABILITIES MAY BE FOUND USING **$RSDST** IF THE NUMBER OF DIFFERENCES IS LESS THAN OR EQUAL TO 30.

ANY PAIR (XI,YI) IS EXCLUDED FROM CONSIDERATION IF XI=YI, AND THE NUMBER OF OBSERVATIONS IS REDUCED ACCORDINGLY. IF THERE ARE TIES AMONG THE ABSOLUTE VALUES OF THE DIFFERENCES, THE AVERAGE RANK OF THE TIED GROUP IS ASSIGNED.

THE LARGE SAMPLE APPROXIMATION TO **TMIN** IS CALCULATED AS FOLLOWS?

```
          T-(N(N+1)/4)
     AT = ------------.
          SQRT(VAR(T))
```

IN GENERAL

```
VAR(T) = [N(N+1)(2N+1)/24] - [SUM(T(I))(T(I)-1)(T(I)+1)/2]
```

WHERE **T** IS **TMIN**, **N** IS THE NUMBER OF DIFFERENCES, **T(I)** IS THE NUMBER OF OBSERVATIONS IN THE ITH TIED GROUP, AND **SUM** INDICATES SUMMING OVER ALL I GROUPS. FOR AN UNITED OBSERVATION **T(I)=1**. WHEN THERE ARE NO TIES THE LARGE SAMPLE APPROXIMATION REDUCES TO

```
      T-(N(N+1)/4)
AT = ---------------------.
      SQRT(N(N+1)(2N+1)/24)
```

FOR LARGE N, AT FOLLOWS THE STANDARD NORMAL DISTRIBUTION. THE PROBABILITY OF FINDING A RANDOM VALUE OF AT GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED AT IS ALSO PRINTED.

REFERENCES

HOLLANDER, M. AND C. A. WOLFE (1973) **NON-PARAMETRIC STATISTICAL METHODS.** WILEY, NEW YORK. PP 27-33.

NOETHER, G. (1971) **INTRODUCTION TO STATISTICS? A FRESH APPROACH.** HOUGHTON-MIFFLIN CO., BOSTON. PP 122-129.

ONE WAY ANALYSIS OF VARIANCE

```
1      8
$ANOVA I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..   INPUT VARIABLES (TREATMENTS)
    ..
    I50
```

OUTPUT

1. THE NUMBERS OF OBSERVATIONS, MEANS, AND STANDARD DEVIATIONS ARE PRINTED FOR EACH VARIABLE, TREATMENT, SPECIFIED.
2. AN ANALYSIS OF VARIANCE TABLE FOR A ONE WAY CLASSIFICATION WITH EQUAL OR UNEQUAL NUMBERS OF REPLICATES IS PRINTED. THE TABLE INCLUDES THE BETWEEN TREATMENTS, WITHIN TREATMENTS, AND TOTAL SUMS OF SQUARES AND DEGREES OF FREEDOM, THE BETWEEN TREATMENTS AND WITHIN TREATMENTS MEAN SQUARES, THE F-RATIO OF THESE TWO MEAN SQUARES, AND THE PROBABILITY, P, OF FINDING A RANDOM VALUE OF -F- GREATER THAN THE OBTAINED VALUE.

USAGE

1. A MAXIMUM OF 50 INPUT VARIABLES MAY BE USED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

CARDS	COMMENTS
1 8	
$ANOVA 5,9,10	A ONE WAY ANALYSIS OF VARIANCE WITH 3 TREATMENTS, VARIABLES 5, 9, AND 10, WILL BE PERFORMED.
$ANOVA ALL	A ONE WAY ANALYSIS OF VARIANCE WILL BE PERFORMED USING ALL INPUT VARIABLES AS TREATMENTS.

IN GENERAL, THE ANALYSIS OF VARIANCE IS A METHOD OF PARTITIONING THE OVERALL VARIATION IN AN EXPERIMENT INTO RECOGNIZED SOURCES OF VARIATION. IN A ONE-WAY DESIGN THE TOTAL VARIATION IS CONSIDERED TO BE EQUAL TO THE SUM OF THE VARIATION BETWEEN TREATMENTS AND THE VARIATION ASSOCIATED WITH EXPERIMENTAL ERROR, (WITHIN TREATMENTS).

IN THE ONE-WAY DESIGN WITH K TREATMENTS AND R(I) REPLICATES IN THE ITH TREATMENT, AN OBSERVATION CAN BE DESCRIBED BY THE MODEL

Y(IJ) = M + T(I) + E(IJ)

FOR I = 1,...,K
J = 1,...,R(I)

WHERE
- Y(IJ) IS THE JTH OBSERVATION IN THE ITH TREATMENT,
- M IS THE OVERALL MEAN,
- T(I) IS THE EFFECT OF THE ITH TREATMENT, AND
- E(IJ) IS A RANDOM ERROR ASSOCIATED WITH THE IJTH OBSERVATION.

IT IS ASSUMED THAT THE SUM OF THE T(I)'S IS ZERO AND THAT THE E(IJ)'S ARE INDEPENDENT AND NORMALLY DISTRIBUTED WITH MEAN 0 AND AND A COMMON VARIANCE. (IT SHOULD BE NOTED THAT THE ASSUMPTIONS ABOUT THE E(IJ)'S ARE MORE STRINGENT IN THIS MODEL THAN IN THE CORRESPONDING MODEL FOR THE KRUSKAL-WALLIS' TEST.)

THE NULL HYPOTHESIS TESTED IS ALL T(I)=0 FOR I=1,...,K. THIS IS EQUIVALENT TO THE HYPOTHESIS THAT THE MEANS OF ALL TREATMENTS ARE EQUAL. THE USUAL ALTERNATE HYPOTHESIS IS NOT ALL T(I)=0. THIS IS EQUIVALENT TO THE STATEMENT THAT AT LEAST ONE PAIR OF TREATMENT MEANS ARE NOT EQUAL.

THE COMPLETE ANALYSIS OF VARIANCE TABLE HAS THE FORM?

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F
TREATMENTS	SSTRT	K-1	MSTRT=SSTRT/K-1	MSTRT/MSE
WITHIN TREATMENTS	SSE	SUM(R(I)-1)	MSE=SSE/SUM(R(I)-1)	
TOTAL	SST	N-1		

WHERE K IS THE NUMBER OF TREATMENTS, R(I) IS THE NUMBER OF REPLICATES IN THE ITH TREATMENT, N IS THE TOTAL NUMBER OF OBSERVATIONS, AND SUM INDICATES SUMMING OVER ALL TREATMENTS. FOR THE DEFINITIONAL AND COMPUTATIONAL FORMULAE OF THE SUMS OF SQUARES, PLEASE SEE THE REFERENCE LISTED BELOW.

THE OUTPUT INCLUDES THE PROBABILITY, P, OF FINDING A RANDOM VALUE OF F LARGER THAN THE OBTAINED F. THIS IS THE SMALLEST SIGNIFICANCE LEVEL AT WHICH THE NULL HYPOTHESIS ALL T(I)=0 MAY BE REJECTED IN FAVOR OF THE ALTERNATE HYPOTHESIS NOT ALL T(I)=0.

REFERENCE

STEELE, R. G. D. AND J. H. TORRIE (1960) **PRINCIPLES AND PROCEDURES OF STATISTICS.** MCGRAW-HILL, NEW YORK. PP 99-119.

DUNCAN'S NEW MULTIPLE RANGE TEST

```
1      8
$DNMRT I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..   INPUT VARIABLES
    ..
    I50
```

OUTPUT

1. THE MEANS OF THE SPECIFIED VARIABLES ARE PRINTED.
2. THE ERROR MEAN SQUARE IS PRINTED.
3. VARIABLES ARE TESTED FOR A SIGNIFICANT DIFFERENCE BETWEEN THE MEANS AT THE .05 LEVEL (ALPHA = .05).

USAGE

1. A MAXIMUM OF 50 VARIABLES MAY BE USED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

```
    CARDS            COMMENTS
    1      8
    $DNMRT 1,3,6     THE MEANS OF VARIABLES 1, 3, AND 5 ARE COMPARED.
    $DNMRT ALL       THE MEANS OF ALL VARIABLES ARE COMPARED.
```

AS MENTIONED PREVIOUSLY, A ONE-WAY ANALYSIS OF VARIANCE MAY BE USED TO TEST THE HYPOTHESIS THAT THE TREATMENT MEANS IN A ONE-WAY DESIGN ARE ALL EQUAL AGAINST THE ALTERNATE HYPOTHESIS THAT AT LEAST ONE PAIR OF TREATMENT MEANS ARE NOT EQUAL. AN INVESTIGATOR MAY BE INTERESTED IN MORE SPECIFIC INFORMATION ABOUT THE TREATMENT MEANS THAN PROVIDED BY THE ANALYSIS OF VARIANCE. IT MAY BE OF INTEREST TO KNOW WHICH TREATMENT MEANS DIFFER. DUNCAN'S NEW MULTIPLE RANGE TEST IS A METHOD OF COMPARING ALL PAIRWISE COMBINATIONS OF TREATMENT MEANS.

CONSIDER AN EXPERIMENT WITH **K** TREATMENTS AND **R** REPLICATES PER TREATMENT. THE TOTAL NUMBER OF OBSERVATIONS IS **N=RK**. THERE ARE THREE STEPS INVOLVED IN DUNCAN'S NEW MULTIPLE RANGE TEST.

I. **SXBAR**, THE STANDARD ERROR OF THE MEAN IS CALCULATED AS

SXBAR=SQRT(ERROR MEAN SQUARE/R)

WHERE **R** IS THE NUMBER OF REPLICATES PER TREATMENT AND THE **ERROR MEAN SQUARE** IS THE EQUIVALENT OF THE WITHIN TREATMENTS MEAN SQUARE IN THE ANALYSIS OF VARIANCE. (NOTE -- WHEN THERE ARE UNEQUAL REPLICATES PER TREATMENT THE HARMONIC MEAN OF THE NUMBERS OF REPLICATES IS USED INSTEAD OF **R**.) **SXBAR** HAS **N-K** DEGREES OF FREEDOM. SIGNIFICANT RANGES AT THE 5¬ LEVEL ARE OBTAINED FOR THE APPROPRIATE DEGREES OF FREEDOM, N-K, AND FOR THE NUMBER OF MEANS TO BE COMPARED, **P=2,...,K**. THESE RANGES ARE CONTAINED IN A TABLE IN THE PROGRAM. THE APPROPRIATE RANGES ARE THEN MULTIPLIED BY **SXBAR** TO GIVE THE CRITICAL VALUE WHICH 2 MEANS **P** STEPS APART ON AN ORDERED SCALE MUST EXCEED TO BE CONSIDERED DIFFERENT AT THE 5¬ LEVEL.

II. THE TREATMENT MEANS ARE RANKED FROM LEAST TO GREATEST.

III. THE DIFFERENCES BETWEEN TREATMENT MEANS ARE COMPARED TO THE CRITICAL VALUES. THE MEANS ARE COMPARED IN THE FOLLOWING ORDER? LARGEST MEAN WITH SMALLEST MEAN, LARGEST WITH SECOND SMALLEST,..., LARGEST WITH SECOND LARGEST, THEN SECOND LARGEST WITH SMALLEST, AND ON THROUGH SECOND SMALLEST WITH SMALLEST. A DIFFERENCE BETWEEN TWO MEANS IS DEEMED SIGNIFICANT IF IT EXCEEDS ITS CRITICAL VALUE AND IS NOT WITHIN A NON-SIGNIFICANT RANGE.

AN EXAMPLE MAY HELP TO CLARIFY THE ABOVE. SUPPOSE WE HAVE FIVE TREATMENT MEANS **A, B, C, D,** AND **E**. IN STEP 1, 4 CRITICAL VALUES, **I, II, III,** AND **IV**, WOULD BE CALCULATED FOR **P** (NUMBER OF STEPS BETWEEN MEANS)= **2, 3, 4,** AND **5**. SUPPOSE THE MEANS ARE RANKED FROM SMALLEST TO LARGEST AS FOLLOWS?

B A E D C.

THE DIFFERENCE **C-B** WOULD BE CONSIDERED FIRST. **C-B** WOULD BE COMPARED TO CRITICAL VALUE **IV** SINCE THE RANGE ENCOMPASSED BY **C** AND **B** CONTAINS 5 MEANS. IF **C-B<IV**, WE WOULD CONCLUDE THAT ALL THE MEANS ARE THE SAME AT THE .05 LEVEL. HOWEVER, IF **C-B>IV**, WE WOULD CONTINUE BY COMPARING **C-A** WITH CRITICAL VALUE **III** AND SO ON UNTIL A NON-SIGNIFICANT RANGE WAS FOUND OR UNTIL THE LAST COMPARISON, **A-B**, WAS MADE.

THE RESULTS ARE SUMMARIZED BY UNDERSCORING MEANS WHICH ARE THE SAME AT THE .05 LEVEL.

REFERENCE
STEELE, R. G. D. AND J. H. TORRIE (1960) **PRINCIPLES AND PROCEDURES OF STATISTICS.** MCGRAW-HILL, NEW YORK. PP 107-109.

NEWMAN - KEUL'S MULTIPLE RANGE TEST

```
1      8
$NKMRT I1,I2,...,I15
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1
I2
..
..    INPUT VARIABLES
..
I15
```

OUTPUT

1. MEANS OF THE SPECIFIED VARIABLES ARE PRINTED.
2. THE ERROR MEAN SQUARE IS COMPUTED.
3. VARIABLES ARE TESTED FOR SIGNIFICANT DIFFERENCES BETWEEN THE MEANS AT THE .05 LEVEL (ALPHA=.05).

USAGE

1. A MAXIMUM OF 15 INPUT VARIABLES MAY BE USED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

CARDS	COMMENTS
1 8	
$NKMRT 1,20,21	THE MEANS OF VARIABLES 1, 20 AND 21 ARE COMPARED.
$NKMRT ALL	THE MEANS OF ALL VARIABLES ARE COMPARED.

THE PROCEDURES INVOLVED IN PERFORMING THE NEWMAN-KEUL'S MULTIPLE RANGE TEST ARE IDENTICAL TO THOSE USED IN DUNCAN'S NEW MULTIPLE RANGE TEST EXCEPT THAT DIFFERENT RANGE VALUES ARE USED. THE RANGE VALUES FOR THE TWO TESTS ARE IDENTICAL WHEN THE NUMBER OF MEANS BEING COMPARED EQUALS 2. WHEN THE NUMBER OF MEANS IN A RANGE EXCEEDS 2, RANGE VALUES FOR THE NEWMAN-KEUL'S TEST ARE LARGER THAN CORRESPONDING RANGE VALUES FOR DUNCAN'S TEST.

REFERENCE

STEELE, R. G. D. AND J. H. TORRIE (1960) **PRINCIPLES AND PROCEDURES OF STATISTICS.** MCGRAW-HILL, NEW YORK. PP 110-111.

THE KRUSKAL-WALLIS STATISTIC

```
1     8
$KRUWL I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1
I2
..
..    INPUT VARIABLES
..
I50
```

OUTPUT

1. THE RANK OF EACH VALUE IN THE DATA SET IS PRINTED. THE SUMS OF THE RANKS FOR EACH VARIABLE ARE GIVEN. TIES ARE GIVEN THEIR AVERAGE RANKING.
2. THE KRUSKAL - WALLIS -H- STATISTIC AND ITS LEVEL OF SIGNIFICANCE ARE PRINTED.

USAGE

1. A MAXIMUM OF 50 VARIABLES MAY BE USED.
2. THE TOTAL NUMBER OF OBSERVATIONS CANNOT EXCEED 5000.

EXAMPLES

```
1     8
CARDS                 COMMENTS
$KRUWL 1,2,6          THE H STATISTIC IS CALCULATED
                      USING VARIABLES 1, 2 AND 6.
$KRUWL ALL            THE H STATISTIC IS CALCULATED
                      USING ALL VARIABLES.
```

THE KRUSKAL-WALLIS TEST IS AN EXTENSION OF WILCOXON↑S RANK SUM TEST FROM 2 INDEPENDENT GROUPS, TREATMENTS OR CLASSIFICATIONS TO K INDEPENDENT GROUPS. THE BASIC MODEL IS

Y(IJ)= M + T(J) + E(IJ)

FOR I=1,...,N(J)
J=1,...,K

WHERE

Y(IJ) IS THE ITH OBSERVATION IN THE GROUP J,
M IS THE OVERALL MEAN,
T(J) IS THE EFFECT OF THE JTH TREATMENT,
E(IJ) IS A RANDOM ERROR ASSOCIATED WITH THE IJTH OBSERVATION, AND
N(J) IS THE NO. OF REPLICATES IN TREATMENT J.

IT IS ASSUMED THAT THE SUM OF THE T(J)↑S IS ZERO, AND THAT THE E(IJ)↑S ARE MUTUALLY INDEPENDENT AND COME FROM THE SAME CONTINUOUS POPULATION.

THE NULL HYPOTHESIS IS HO? T(1)=T(2)=...=T(J), AGAINST THE ALTERNATIVE THAT NOT ALL THE T(J)↑S ARE EQUAL.

ALL N=N(1)+N(2)+---+N(K) OBSERVATIONS ARE ORDERED FROM LEAST TO GREATEST AND RANKS 1 TO N ARE ASSIGNED. THE SUM OF THE RANKS IS THEN COMPUTED FOR EACH TREATMENT GROUP. IF R(IJ) IS THE RANK OF Y(IJ), THEN THE SUM OF THE RANKS FOR TREATMENT J IS R(J)= SUM(R(IJ)), WHERE SUM INDICATES SUMMING OVER I FROM 1 TO N(J). THE AVERAGE RANK FOR TREATMENT J IS MR(J)=R(J)/N(J). THE OVERALL AVERAGE OF THE RANKS IS MR=(N+1)/2.

THE KRUSKAL-WALLIS STATISTIC IS DEFINED AS

H = [12/N(N+1)][SUM(N(J))(MR(J) - MR)↑2)],

WHERE SUM INDICATES SUMMING OVER J FROM 1 TO K. THIS FORMULA CAN BE CONVERTED TO THE COMPUTATIONAL FORMULA

H = [(12/N(N+1))(SUM((R(J)↑2)/N(J))]-3(N+1).

FOR LARGE N(J), H FOLLOWS AN APPROXIMATE CHI-SQUARE DISTRIBUTION WITH K-1 DEGREES OF FREEDOM. THE PROBABILITY OF FINDING A LARGER H BY CHANCE ALONE IS GIVEN AS THE LEVEL OF SIGNIGICANCE. THIS IS THE LOWEST SIGNIFICANCE LEVEL AT WHICH THE NULL HYPOTHESIS MAY BE REJECTED IN FAVOR OF THE ALTERNATE HYPOTHESIS.

WHEN THERE ARE TIED OBSERVATIONS, THE AVERAGE OF THE RANKS OF THE TIED VALUES ARE ASSIGNED AND THE STATISTIC

H* = H/[1-((SUM(S(L)↑3-S(L))/(N↑3-N))]

IS USED INSTEAD OF H. IN THE ABOVE FORMULA, SUM INDICATES SUMMING OVER L FROM 1 TO G WHERE G IS THE NUMBER OF TIED GROUPS, AND S(L) IS THE NUMBER OF TIED VALUES IN GROUP L. FOR UNTIED OBSERVATIONS S(L)=1. FOR DATA WITH NO TIES H*=H.

FOR SMALL N(J) THE USER IS REFERRED TO TABLES OF THE KRUSKAL-WALLIS, H, STATISTIC TO DETERMINE SIGNIFICANCE OF H. THESE TABLES ARE TOO LENGTHY TO BE PRESENTED HERE. ONE SOURCE IS TABLE A.7 IN HOLLANDER AND WOLFE (1973).

REFERENCES

HOLLANDER, M. AND D. A. WOLFE (1973). **NON-PARAMETRIC STATISTICAL METHODS.** WILEY, NEW YORK. PP 115-119.

NOETHER, G. (1971). **INTRODUCTION TO STATISTICS? A FRESH APPROACH.** HOUGHTON-MIFFLIN CO., BOSTON. PP143-146.

MULTIPLE COMPARISONS BASED ON THE KRUSKAL-WALLIS H STATISTIC

```
1     8
$KRUWC I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..   INPUT VARIABLES
    ..
    I50
```

OUTPUT

1. ALL OUTPUT GIVEN BY $KRUWL.
2. THE MULTIPLE COMPARISON PROCEDURE INCLUDES A LIST OF THE TREATMENT, VARIABLE, PAIRS COMPARED, THE ABSOLUTE DIFFERENCE BETWEEN THE AVERAGE RANKS OF TREATMENT PAIRS, THE TEST STATISTICS, Z(I,J)'S, AND THE LEVELS OF SIGNIFICANCE OF THE Z(I,J)'S.

USAGE

1. A MAXIMUM OF 50 VARIABLES MAY BE USED.
2. THE TOTAL NUMBER OF OBSERVATIONS CANNOT EXCEED 5000.

EXAMPLE

CARD	COMMENTS
1 8	
$KRUWC 1,3,6	THE H STATISTIC IS CALCULATED USING VARIABLES 1, 3 AND 6. THE DIFFERENCES OF THE AVERAGE RANKS OF THE VARIABLE PAIRS (1,3), (1,6) AND (3,6) ARE TESTED.

THIS PROGRAM CALCULATES THE **KRUSKAL-WALLIS H STATISTIC** AS IN **$KRUWL**, AND PERFORMS ALL PAIRWISE COMPARISONS BETWEEN THE AVERAGE RANKS OF THE K TREATMENTS.

IF THERE ARE K TREATMENTS, K(K-1)/2 COMPARISONS ARE PERFORMED.

THE STATISTICS CALCULATED FOR ALL I,J, AND I≠J ARE

$$Z(IJ) = \frac{ABS[RBAR(I) - RBAR(J)]}{SQRT[N(N-1)[1/N(I) + 1/N(J)]/12]}$$

WHERE

RBAR(I) AND RBAR(J) ARE THE AVERAGE RANKS OF THE ITH AND JTH TREATMENTS

N = TOTAL NUMBER OF OBSERVATIONS

N(I) AND N(J) = NUMBER OF OBSERVATIONS IN ITH AND JTH TREATMENTS RESPECTIVELY.

Z(IJ) FOLLOWS A STANDARD NORMAL DISTRIBUTION. THE LEVELS OF SIGNIFICANCE GIVEN ON THE OUTPUT ARE THE UPPER TAIL PROBABILITIES FOR EACH Z(IJ). TO MAKE A DECISION CONCERNING ANY GIVEN PAIR, CHOOSE A DESIRED OVERALL SIGNIFICANCE LEVEL, ALPHA. THIS WILL INSURE THAT THE PROBABILITY WILL BE AT MOST ALPHA OF DECLARING THAT TWO OR MORE TREATMENTS DIFFER WHEN IN FACT ALL K TREATMENTS ARE IDENTICAL. THEN COMPUTE C = ALPHA/K(K-1). C CAN THEN BE COMPARED WITH THE LEVELS OF SIGNIFICANCE GIVEN FOR EACH Z(IJ). IF THE LEVEL OF SIGNIFICANCE FOR A PARTICULAR Z(IJ) IS LESS THAN OR EQUAL TO C, WE MAY DECIDE THAT THE AVERAGE RANKS FOR THE ITH AND JTH TREATMENT DIFFER AT THE CHOSEN OVERALL SIGNIFICANCE LEVEL ALPHA.

REFERENCES?

1. NOETHER, G. (1971) **INTRODUCTION TO STATISTICS? A FRESH APPROACH.** HOUGHTON-MIFFLIN CO., BOSTON. PP 147-148.

PEARSON'S CORRELATION

```
1      8
$CORR  I1,I2,...,I25
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1
I2
..
..    INPUT VARIABLES
..
I25
```

OUTPUT

1. PEARSON'S CORRELATION COEFFICIENTS BETWEEN ALL PAIRWISE COMBINATIONS OF INPUT VARIABLES AND CORRESPONDING NUMBERS OF OBSERVATIONS ARE PRINTED.
2. THE FORM OF THE OUTPUT IS (OBSERVATIONS, CORRELATION).

USAGE

1. A MAXIMUM OF 25 VARIABLES MAY BE USED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

CARDS		COMMENTS
1	8	
$CORR	1,4,6,	CORRELATIONS OF VARIABLE PAIRS (1,4), (1,6) AND (4,6) ARE COMPUTED.
$CORR	ALL	CORRELATIONS OF ALL PAIRS OF VARIABLES ARE CALCULATED.

A **CORRELATION COEFFICIENT** IS A MEASURE OF THE DEGREE OF ASSOCIATION BETWEEN TWO VARIABLES. THERE ARE MANY TYPES OF CORRELATION COEFFICIENTS. THE ONE CALCULATED BY THIS PROGRAM IS **THE PRODUCT MOMENT CORRELATION COEFFICIENT, R.** SPECIFICALLY, R IS A MEASURE OF THE LINEAR ASSOCIATION BETWEEN TWO VARIABLES. THE VALUES OF R CAN RANGE FROM -1 (SMALL VALUES OF ONE VARIABLE ARE LINEARLY ASSOCIATED WITH LARGE VALUES OF THE OTHER VARIABLE) TO +1 (VALUES OF THE TWO VARIABLES TEND TO INCREASE TOGETHER). A VALUE OF 0 INDICATES NO LINEAR ASSOCIATION BETWEEN THE VARIABLES.

THE **PRODUCT MOMENT CORRELATION** IS DEFINED AS

$$R = \frac{SUM[[X(I)-XBAR][Y(I)-YBAR]]}{SQRT[[SUM(X(I)-XBAR)\uparrow 2][SUM(Y(I)-YBAR)\uparrow 2]]}$$

WHERE

SUM INDICATES SUMMING OVER VALUES OF I = 1,2,...,N

↑ IS EXPONENTIATION

X(I) = THE ITH VALUE OF VARIABLE X

Y(I) = THE ITH VALUE OF VARIABLE Y

XBAR AND YBAR ARE THE SAMPLE MEANS OF X AND Y RESPECTIVELY

THE COMPUTING ALGORITHM USED IS

$$R = \frac{SUM\ AB - (SUM\ A)\ (SUM\ B)/N}{SQRT[[SUM(A\uparrow 2)-((SUM\ A)\uparrow 2)/N][SUM(B\uparrow 2)-((SUM\ B)\uparrow 2)/N]]}$$

WHERE

A IS X(I)

B IS Y(I)

N = NUMBERS OF X AND Y VALUES

FOR TESTS OF HYPOTHESES ABOUT **RHO**, THE POPULATION CORRELATION COEFFICIENT (ESTIMATED BY R), SEE STEELE AND TORRIE (1960).

REFERENCES?

STEELE, R.G.D. AND J.H. TORRIE (1960) **PRINCIPLES AND PROCEDURES OF STATISTICS.** MCGRAW-HILL, NEW YORK. PP 183-193.

KENDALL'S RANK CORRELATION

```
1      8
$RNKCR I1,I2,...,I25
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..    INPUT VARIABLES
    ..
    I25
```

OUTPUT

1. THE RANK OF EACH VALUE WITHIN A VARIABLE AND THE VALUES OF EACH VARIABLE IN A VARIABLE PAIR ARE PRINTED.
2. THE -S- STATISTIC, KENDALL'S CORRELATION COEFFICIENT, (TAU), THE LARGE SAMPLE APPROXIMATION TO -S-, AND THE SIGNIFICANCE LEVEL OF THE LARGE SAMPLE APPROXIMATION ARE GIVEN FOR ALL PAIRWISE COMBINATIONS OF THE VARIABLES REQUESTED.

USAGE

1. A MAXIMUM OF 25 OBSERVATIONS MAY BE USED.
2. THE MAXIMUM NUMBER OF OBSERVATIONS PER VARIABLE IS 500.

EXAMPLES

CARDS	COMMENTS
1 8	
$RNKCR 1,4,6	RANK CORRELAITON FOR VARIABLE PAIRS (1,4), (1,6) AND (4,6)
$RNKCR ALL	RANK CORRELATION FOR ALL VARIABLE PAIRS

KENDALL'S RANK CORRELATION IS A METHOD OF DISCERNING THE ASSOCIATION, IF ANY, BETWEEN TWO VARIABLES.

SUPPOSE WE HAVE TWO VARIABLES WITH N OBSERVATIONS EACH? (X(1), Y(1)), (X(2),Y(2)),...,(X(N),Y(N)). THE HYPOTHESIS OF INTEREST IS HO? X AND Y ARE INDEPENDENT (UNASSOCIATED). AN INTUITIVE APPROACH TO DISCERNING THE RELATIONSHIP BETWEEN X AND Y IS TO COUNT HOW OFTEN THE RELATIVE RANKINGS WITHIN EACH VARIABLE MOVE IN THE SAME DIRECTION AND HOW OFTEN THEY MOVE IN OPPOSITE DIRECTIONS.

IN PARTICULAR FOR I AND J=1,2,...,N AND I≠J, IF THE PRODUCT [X(I)-X(J)][Y(I)-Y(J)] IS GREATER THAN 0, (I.E., THE OBSERVATIONS ARE CHANGING IN THE SAME DIRECTION) WE COUNT +1. CONVERSELY, IF THE PRODUCT IS NEGATIVE, WE COUNT -1. IF THE PRODUCT IS 0, WE COUNT 0, BECAUSE AT LEAST ONE OF THE VARIABLES HAD NO CHANGE BETWEEN OBSERVATIONS I AND J.

THE SUM OF THE 1'S, -1'S, AND 0'S IS THE STATISTIC **S**. IF ALL N OBSERVATIONS OF X AND Y MOVED IN THE SAME DIRECTION, **S** WOULD EQUAL N(N-1)/2, WHILE IF ALL OBSERVATIONS MOVED IN OPPOSITE DIRECTIONS, S WOULD EQUAL -N(N-1)/2. IT WOULD BE HELPFUL TO HAVE A STATISTIC WHICH GAVE THE STRENGTH OF ASSOCIATION BETWEEN THE TWO VARIABLES ON A SCALE FROM -1 (EVERY INCREASE IN ONE VARIABLE IS ACCOMPANIED BY A DECREASE IN THE OTHER VARIABLE) TO +1 (THE VARIABLES INCREASE TOGETHER). SUCH A STATISTIC IS **TAU** WHERE **TAU** IS THE OBTAINED VALUE OF S DIVIDED BY THE **MAXIMUM OF S.** THE MAXIMUM OF S IS N(N-1)/2 SO

$$TAU = \frac{S}{N(N-1)/2} = \frac{2S}{N(N-1)}$$

THE SIGNIFICANCE OF TAU CAN BE EVALUATED FOR LARGE SAMPLES IN THE **FOLLOWING WAY?** THE LARGE SAMPLE APPROXIMATION IS

$$S* = \frac{S}{SQRT[N(N-1)(2N+5)/18]}$$

FOR LARGE N, THE LARGE SAMPLE APPROXIMATION FOLLOWS THE STANDARD NORMAL DISTRIBUTION. THE LEVEL OF SIGNIFICANCE PRINTED IS TWICE THE UPPER TAIL PROBABILITY OF THE ABSOLUTE VALUE OF **S.**

FOR N GREATER THAN 8, THE NORMAL APPROXIMATION IS ADEQUATE. FOR N LESS THAN OR EQUAL TO 8, THE READER IS REFERRED TO TABLE A.21 IN HOLLANDER AND WOLFE (1973).

REFERENCES?

1. HOLLANDER, M. AND D.A. WOLFE (1973) NONPARAMETRIC STATISTICAL METHODS. WILEY, NEW YORK. PP 185-193, 384-393.

2. NOETHER, G. (1971) INTRODUCTION TO STATISTICS? A FRESH APPROACH. HOUGHTON-MIFFLIN CO., BOSTON. PP 155-162.

CHI-SQUARE

```
1     8
$CHISQ I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1
I2
..
..    INPUT VARIABLES (COLUMNS)
..
I50
```

OUTPUT

1. A CHI-SQUARE STATISTIC IS CALCULATED FOR A TWO-WAY CONTINGENCY TABLE IN WHICH THE INPUT VARIABLES FORM THE COLUMNS AND THE CASES FORM THE ROWS.
2. DEGREES OF FREEDOM AND SIGNIFICANCE LEVEL OF THE CHI-SQUARE VALUE AND THE CONTINGENCY COEFFICIENT ARE GIVEN.

USAGE

1. A MAXIMUM OF 50 VARIABLES MAY BE USED.
2. THERE IS A LIMIT OF 100 CASES.
3. THE PRODUCT OF THE NUMBER OF VARIABLES TIMES THE NUMBER OF CASES MAY NOT EXCEED 1400.

EXAMPLES

CARDS	COMMENTS
1 8	
$CHISQ 1,2,3	A CHI-SQUARE TEST WILL BE CALCULATED USING VARIABLE 1, 2 AND 3 AS COLUMNS AND ALL CASES AS ROWS.
$CHISQ ALL	ALL VARIARIABLES AND CASES WILL BE USED TO PERFORM A CHI-SQUARE TEST.

THE CHI-SQUARE STATISTIC, **CHI-SQUARE = SUM[((OBS-EXP)↑2)/EXP]**, MAY BE USED TO TEST A NUMBER OF HYPOTHESES. THIS PROGRAM COMPUTES A CHI-SQUARE STATISTIC TO TEST THE HYPOTHESIS OF INDEPENDENCE OF THE 2 VARIABLES OF CLASSIFICATION IN AN **RXC** CONTINGENCY TABLE, WHERE **R** IS THE NUMBER OF ROWS, (DISCRETE CATEGORIES OF ONE VARIABLE), AND **C** IS THE NUMBER OF COLUMNS, (DISCRETE CATEGORIES OF THE OTHER VARIABLE). THE DATA ARE THE NUMBERS OF OBJECTS OR SUBJECTS FALLING INTO EACH OF THE **RXC** CELLS.

AN RXC CONTINGENCY TABLE

```
                      VARIABLE  A

          *********************************
          *  1  *  2  *  .  .  .  .  *  C  *
     *************************************
     *    *     *     *              *     *
     *1   * N11 * N12 *  .  .  .  .  * N1C * N1.
  V  *    *     *     *              *     *
  A  ****--------------------------------*
  R  *    *     *     *              *     *
  I  *2   * N21 * N22 *  .  .  .  .  * N2C * N2.
  A  *    *     *     *              *     *
  B  ****--------------------------------*
  L  *.   *  .  *  .  *  .  .  .  .  *  .  *
  E  *.   *  .  *  .  *  .  .  .  .  *  .  *
     *.   *  .  *  .  *  .  .  .  .  *  .  *
  B  ****--------------------------------*
     *    *     *     *              *     *
     *R   * NR1 * NR2 *  .  .  .  .  * NRC * NR.
     *    *     *     *              *     *
     *************************************
            N.1   N.2                  N.C   N..
```

IF THE TWO VARIABLES OF CLASSIFICATION ARE INDEPENDENT, THE PROBABILITY OF AN OBSERVATION FALLING IN THE IJTH CELL, (I=1,...R, J=1,...,C), WILL BE EQUAL TO THE PROBABILITY OF THE OBSERVATION FALLING IN THE ITH ROW TIMES THE PROBABILITY OF THE OBSERVATION FALLING IN THE JTH COLUMN. EXPECTED NUMBERS OF OBSERVATIONS IN EACH CELL ARE THEN CALCULATED ASSUMING THE ROWS AND COLUMNS ARE INDEPENDENT. SPECIFICALLY, THIS PROGRAM CALCULATES THE EXPECTED NUMBERS FOR EACH IJTH CELL AS: **EXP(IJ)=(NI.)(N.J)/N..**, WHERE **NI.** IS THE TOTAL NUMBER OF OBSERVATIONS IN THE ITH ROW, **N.J** IS THE TOTAL NUMBER OF OBSERVATIONS IN THE JTH COLUMN, AND **N..** IS THE TOTAL NUMBER OF OBSERVATIONS IN THE EXPERIMENT. IT MAKES SENSE INTUITIVELY THAT THE LARGER THE DEVIATIONS OF THE OBSERVED AND EXPECTED VALUES, THE LESS LIKELY IT IS THAT THE ROWS AND COLUMNS ARE INDEPENDENT. THE CHI-SQUARE STATISTIC PROVIDES A METHOD OF DECIDING WHAT THE PROBABILITY IS THAT THE ROWS AND COLUMNS ARE INDEPENDENT, BASED ON THE MAGNITUDE OF THE DEVIATIONS BETWEEN THE OBSERVED AND EXPECTED NUMERS IN EACH CELL. THE FOLLOWING STATISTIC IS COMPUTED.

CHI-SQUARE=SUM[((OBS(IJ)-EXP(IJ))↑2)/EXP(IJ)],

WHERE **SUM** INDICATES SUMMING OVER I AND J FOR **I=1,...,R**, AND **J=1,...,C**. **CHI-SQUARE** FOLLOWS AN APPROXIMATE CHI-SQUARE DISTRIBUTION WITH **(R-1)(C-1)** DEGREES OF FREEDOM. THE PROBABILITY OF FINDING A GREATER

CHI-SQUARE VALUE BY CHANCE ALONE IS GIVEN. THIS IS THE LOWEST SIGNIFICANCE LEVEL AT WHICH THE NULL HYPOTHESIS OF INDEPENDENCE MAY BE REJECTED.

FOR 2X2 TABLES, (R=2,C=2), THE CHI-SQUARE STATISTIC IS CALCULATED WITH A CORRECTION FOR CONTINUITY AND WITHOUT SUCH CORRECTION. THE CHI-SQUARE CORRECTED FOR CONTINUITY IS CALCULATED AS

```
             [[ABS((N11)(N22)-(N12)(N21))-N../2]↑2]N..
CHI-SQUARE = -----------------------------------------
                  (N1.)(N2.)(N.1)(N.2)
```

WITH 1 DEGREE OF FREEDOM. THE CHI-SQUARE STATISTIC FOR THE 2X2 TABLE WITHOUT CORRECTION FOR CONTINUITY IS CALCULATED AS ABOVE BUT WITHOUT THE TERM N../2. GENERALLY, THE CORRECTED CHI-SQUARE GIVES A BETTER APPROXIMATION TO THE CHI-SQUARE DISTRIBUTION. HOWEVER, SOME PREFER NOT TO USE IT.

REFERENCE

STEELE, R. G. D. AND J.H. TORRIE (1960). **PRINCIPLES AND PROCEDURES OF STATISTICS.** MCGRAW-HILL, NEW YORK. PP 366-375.

LEAST SQUARES REGRESSION

```
1      8
$LSQR  I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1 - DEPENDENT VARIABLES
    I2 -
    ..
    ..   INDENDENT VARIABLES
    ..
    I50 -
```

OUTPUT

FOR EACH COMBINATION OF THE DEPENDENT VARIABLE WITH AN INDEPENDENT VARIABLE -

1. INTERCEPT AND SLOPE FOR THE LEAST SQUARES LINE ARE GIVEN.
2. T-STATISTIC WITH DEGREES OF FREEDOM AND SIGNIFICANCE LEVEL ARE COMPUTED TO TEST THE HYPOTHESIS THAT THE SLOPE IS ZERO.
3. THE NUMBER OF OBSERVATIONS IS GIVEN.
4. THE STANDARD ERROR OF THE SLOPE ESTIMATE, THE CORRELATION COEFFICIENT, AND THE COEFFICIENT OF DETERMINATION ARE GIVEN.
5. THE F-STATISTIC, DEGREES OF FREEDOM, AND SIGNIFICANCE LEVEL FOR TESTING THE OVERALL SIGNIFICANCE OF REGRESSION ARE PRINTED, AND
6. MEANS AND STANDARD DEVIATIONS ARE PRINTED.

USAGE

1. A MAXIMUM OF ONE DEPENDENT AND 49 INDEPENDENT VARIABLES ARE ALLOWED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLE

CARD	COMMENTS
1 8	
$LSQR 1,3,6	TWO LEAST SQUARES ANALYSES WILL BE PERFORMED- (DEPENDENT VARIABLE-1 WITH INDEPENDENT VARIABLE-3) AND (DEPENDENT VARIABLE-1 WITH INDEPENDENT VARIABLE-6).

SIMPLE LINEAR REGRESSION IS A METHOD OF FITTING A STRAIGHT LINE TO A COLLECTION OF POINTS (Y(I),X(I)). THE MODEL IS

Y(I)= **B0** + **B1**X(I) + E(I)

WHERE Y(I) AND X(I) ARE THE ITH OBSERVATIONS OF Y AND X (I=1,...,N),
B0 = THE UNKNOWN INTERCEPT PARAMETER,
B1 = THE UNKNOWN SLOPE PARAMETER, AND
E(I) = AN UNKNOWN INCREMENT BY WHICH ANY Y(I) MAY VARY FROM THE TRUE LINE Y=**B0** + **B1**X.

LEAST SQUARES IS A METHOD OF ESTIMATING **B0** AND **B1** WITH B0 AND B1 (ESTIMATES ARE NOT IN BOLDFACE) SO THAT

SUM (E(I)↑2)= SUM[[Y(I)-B0-B1X(I)]↑2]

IS A MINIMUM. THAT IS, SO THAT THE SUM OF THE SQUARED DEVIATIONS OF THE PREDICTED Y↑S, YHAT(I)= B0 + B1X(I), AND THE OBSERVED Y↑S IS A MINIMUM.

THE ESTIMATES OF **B0** AND **B1** ARE

```
      SUM[(X(I)-MX)(Y(I)-MY)]    SUM[(X(I))(Y(I))]-[(SUMX(I))(SUMY(I))]/N
B1 =------------------------ = -------------------------------------------
        SUM[(X(I)-MX)↑2]                SUM(X(I)↑2)-[(SUMX(I))↑2]/N
```

AND

B0 = MY - B1MX

WHERE MY AND MX ARE THE MEANS OF Y AND X, AND SUM INDICATES SUMMING OVER I FOR I=1,...,N.

ANOTHER OBJECT OF LINEAR REGRESSION IS TO TEST HYPOTHESES ABOUT THE RELATIONSHIP BETWEEN THE 2 VARIABLES OF INTEREST. THE MOST FREQUENTLY STATED NULL HYPOTHESIS IS THAT THERE IS NO RELATIONSHIP BETWEEN THE VARIABLES: HO? **B1**=0 AGAINST POSSIBLE ALTERNATIVES, **B1**≠0, **B1**>0, **B1**<0. THE TEST STATISTIC IS

T = B1/STANDARD ERROR OF B1.

T HAS N-2 DEGREES OF FREEDOM. THE COMPUTATIONAL FORMULA FOR THE STANDARD ERROR OF B1 IS RATHER LENGTHY AND WILL NOT BE PRESENTED HERE, BUT MAY BE FOUND IN THE REFERENCE GIVEN.

ANOTHER EQUIVALENT TEST IS THE F TEST FOR SIGNIFICANCE OF REGRESSION. THE COMPUTATION OF THE F-STATISTIC IS MOST INTELLIGIBLE PRESENTED IN TABULAR FORM. THE SOURCES OF VARIATION IN THE REGRESSION MODEL MAY BE PARTIONED AS TOTAL VARIATION= VARIATION DUE TO REGRESSION + VARIATION ABOUT REGRESSION. THE VARIATION ASSOCIATED WITH EACH SOURCE ARE DEFINED AS IN THE TABLE BELOW. COMPUTATIONAL FORMULA OF SST AND SSR MAY BE FOUND IN THE REFERENCE GIVEN.

```
*****************************************************************
SOURCE       * DF* SUMS OF SQUARES                    * MEAN SQUARES
-----------------------------------------------------------------
TOTAL        *N-1* SST= SUM[(Y(I)-MY)↑2]              *
             *   *                                    *
DUE TO       *   * SSR=                               *
REGRESSION* 1 *  B1[SUM[(X(I)-MX)(Y(I)-MY)]]          * MSR=SSR
             *   *                                    *
ABOUT        *   *                                    *
REGRESSION*N-2* SSE=SUM[(Y(I)-YHAT)↑2]=SST-SSR* MSE=SSE/N-2
*****************************************************************
```

THE TEST STATISTIC IS **F=MSR/MSE** WITH 1 AND N-2 DEGREES OF FREEDOM. NOTE THAT **T↑2=F**. THE SIGNIFICANCE LEVELS OF **T** AND **F** ARE GIVEN.

OTHER QUANTITIES CALCULATED ARE THE CORRELATION COEFFICIENT R, AND THE COEFFICIENT OF DETERMINATION, R↑2=SSR/SST. R↑2 IS A MEASURE OF THE PROPORTION OF THE TOTAL VARIATION ABOUT THE MEAN, MY, EXPLAINED BY THE REGRESSION OF Y ON X. NATURALLY, THE LARGER R↑2 THE BETTER THE FIT IS.

REFERENCE

DRAPER, N. R. AND H. SMITH (1966). **APPLIED REGRESSION ANALYSIS.** JOHN WILEY AND SONS, INC., NEW YORK. PP 1-34.

MULTIPLE REGRESSION

```
1       8
$MREG   I1,I2,...,I25
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1 - DEPENDENT VARIABLE
    I2 -
    ..
    ..    INDEPENDENT VARIABLES
    ..
    I25 -
```

OUTPUT

1. THE MEANS, AND STANDARD DEVIATIONS OF ALL VARIABLES ARE PRINTED.
2. CORRELATIONS BETWEEN ALL VARIABLES ARE GIVEN.
3. INTERCEPT, REGRESSION COEFFICIENTS, STANDARD ERRORS OF REGRESSION COEFFICIENTS, T-VALUES, DEGREES OF FREEDOM AND SIGNIFICANCE LEVELS OF THE T-VALUES FOR TESTING HYPOTHESES THAT THE REGRESSION COEFFICIENTS ARE ZERO ARE INCLUDED.
4. MULTIPLE CORRELATION AND DETERMINATION ARE GIVEN.
5. THE STANDARD ERROR OF THE ESTIMATE IS PRINTED.
6. AN ANALYSIS OF VARIANCE FOR THE REGRESSION IS GIVEN. THIS INCLUDES THE SUMS OF SQUARES ATTRIBUTABLE TO REGRESSION, ABOUT THE REGRESSION, AND OF THE TOTAL. DEGREES OF FREEDOM, MEAN SQUARES, THE F-STATISTIC AND ITS LEVEL OF SIGNIFICANCE ARE PRINTED.

USAGE

1. ONE DEPENDENT VARIABLE MAY BE USED.
2. NO MORE THAN 24 INDEPENDENT VARIABLES MAY BE USED.
3. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

CARDS	COMMENTS
1 8	
$MREG ALL	A MULTIPLE REGRESSION ANALYSIS USING VAR. 1 AS THE DEPENDENT VARIABLE AND REMAINING VARIABLES AS THE INDEPENDENT VARIABLES WILL BE PERFORMED.
$MREG 3,1,2	A MULTIPLE REGRESSION ANALYSIS WILL BE PERFORMED USING VAR. 3 AS THE DEPENDENT VARIABLE AND VARIABLES 1 AND 2 AS THE INDEPENDENT VARIABLES.

IN MULTIPLE REGRESSION, THE MODEL IS
Y(I)= B0 + B1*X1(I) + B2*X2(I) + ... + BK*XK(I) + E(I),

WHERE Y(I) IS THE ITH MEASUREMENT OF THE DEPENDENT VARIABLE Y, (I=1,...,N),
X1(I), X2(I),...,XK(I) ARE THE ITH MEASUREMENTS OF THE INDEPENDENT VARIABLES, X1, X2,...,XK (I=1,....N),
B0, B1, B2,...,BK ARE THE UNKNOWN COEFFICIENTS ASSOCIATED WITH THE INTERCEPT AND THE K INDEPENDENT VARIABLES, AND
E(I) = ERROR ASSOCIATED WITH THE ITH MEASUREMENT.

ONE OBJECT OF THIS PROGRAM IS TO ESTIMATE B0, B1,...,BK SO THAT THE SUM(E(I)↑2) IS A MINIMUM. ANOTHER OBJECT IS TO DECIDE IF THE OVERALL REGRESSION OF Y ON X1, X2,...,XK IS SIGNIFICANT (I.E., ACCOUNTS FOR OBSERVED VARIATION IN Y). IN THIS REGARD, A TEST IS PROVIDED OF THE HYPOTHESIS HO? B1=B2=...=BK=0, AGAINST THE ALTERNATE HYPOTHESIS, NOT ALL BI=0 FOR I=1,...,K. FURTHER TESTS ARE PROVIDED FOR HYPOTHESES OF THE FORM HO? BI=0, I=1,...,K.

THE FOLLOWING STATISTICS ARE OUTPUT.

1. THE MEANS AND STANDARD DEVIATIONS OF Y AND X1,...,XK ARE PRINTED.

2. THE CORRELATION MATRIX OF THE DEPENDENT AND INDEPENDENT VARIABLES IS GIVEN. THIS GIVES MEASURES OF ASSOCIATION BETWEEN THE DEPENDENT VARIABLE AND EACH INDEPENDENT VARIABLE AND AMONG THE INDEPENDENT VARIABLES.

3. THE ESTIMATES OF B1,...,BK AND THE STANDARD ERRORS OF THE ESTIMATES ARE PRINTED. T STATISTICS ARE PRINTED FOR TESTING EACH HYPOTHESIS HO? BI=0, I=1,...,K.

T= BI/STANDARD ERROR(BI) WITH N-K DEGREES OF FREEDOM.

THE LEVEL OF SIGNIFICANCE IS GIVEN FOR THE TEST OF EACH NULL HYPOTHESIS AGAINST THE TWO-SIDED ALTERNATIVE BI≠0. FOR TESTS AGAINST ONE-SIDED ALTERNATIVES, BI>0 OR BI<0, USE THE PRINTED LEVEL OF SIGNIFICANCE DIVIDED BY 2.

4. THE ESTIMATE OF B0 IS OUTPUT AS THE INTERCEPT.

5. MULTIPLE CORRELATION, R, IS A MEASURE OF THE COMBINED EFFECT OF THE INDEPENDENT VARIABLES ON THE DEPENDENT VARIABLES.

6. THE COEFFICIENT OF MULTIPLE DETERMINATION, R↑2, IS A MEASURE OF THE PROPORTION OF THE TOTAL VARIATION ABOUT THE MEAN OF Y EXPLAINED BY THE REGRESSION OF Y ON X1,...,XK. IT CAN BE THOUGHT OF AS A MEASURE OF THE USEFULNESS OF THE INDEPENDENT VARIABLES IN ACCOUNTING FOR VARIATION IN THE Y(I)↑S.

7. THE STANDARD ERROR OF THE ESTIMATE IS AN ESTIMATE OF THE PRECISION OF THE REGRESSION. IT IS EQUAL TO THE SQRT(MEAN SQUARE ABOUT THE REGRESSION).

8. AN ANALYSIS OF VARIANCE TABLE FOR THE OVERALL REGRESSION IS INCLUDED. THE SOURCES OF VARIATION AND ASSOCIATED DEGREES OF FREEDOM, SUMS OF SQUARES AND MEAN SQUARES ARE GIVEN. AN F-STATISTIC IS

CALCULATED AS THE RATIO OF THE MEAN SQUARE DUE TO REGRESSION/MEAN SQUARE ABOUT THE REGRESSION. THE DEGREES OF FREEDOM OF THE CALCULATED F ARE K,N-K (K=NO. OF PARAMETERS, N= NO. OF OBSERVATIONS). THE F-STATISTIC MAY BE USED TO TEST THE HYPOTHESIS, HO? B1=B2=...=BK AGAINST THE ALTERNATIVE, NOT ALL BI=0, I=1,...,K. THE LEVEL OF SIGNIFICANCE IS THE PROBABILITY OF OBTAINING A LARGER F WITH K,N-K D.F. BY CHANCE ALONE.

REFERENCE

DRAPER, N. R. AND H. SMITH (1966). **APPLIED REGRESSION ANALYSIS.** JOHN WILEY AND SONS. INC., NEW YORK. PP 44-85.

STEP-WISE MULTIPLE REGRESSION

```
1      8
$SWREG I1,I2,....,I25
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1  - DEPENDENT VARIABLE
I2
..
..  - INDEPENDENT VARIABLES
..
I25
```

OUTPUT

THE OUTPUT INCLUDES

1. MEANS AND STANDARD DEVIATIONS OF ALL VARIABLES;
2. CORRELATION MARTIX;
3. SUM OF SQUARES REDUCED IN EACH STEP;
4. CUMULATIVE SUM OF SQUARES REDUCED;
5. MUTIPLE CORRELATION COEFFICIENTS FOR EACH STEP;
6. F-VALUE, DEGREES OF FREEDOM, AND SIGNIFICANCE LEVEL OF F-VALUE FOR TESTING SIGNIFICANCE OF OVERALL REGRESSION AT EACH STEP;
7. STANDARD ERROR OF ESTIMATE, AND
8. AT EACH STEP FOR EACH INDEPENDENT VARIABLE ENTERED, REGRESSION COEFFICIENTS, STANDARD ERROR OF REGRESSION COEFFICIENTS, AND T-VALUES WITH DEGREES OF FREEDOM AND SIGNIFICANCE LEVEL FOR TESTING HYPOTHESES THAT REGRESSION COEFFICIENTS ARE ZERO.

USAGE

1. ONE DEPENDENT VARIABLE MAY BE USED.
2. NO MORE THAN 24 INDEPENDENT VARIABLES MAY BE USED.
3. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLES

CARDS 1 8	COMMENTS
$SWREG ALL	STEPWISE REGRESSION WILL BE PERFORMED WITH VAR. 1 AS THE DEPENDENT VARIABLE, AND REMAINING VARIABLES AS INDEPENDENT VARIABLES.
$SWREG 4,3,2,1	STEPWISE REGRESSION WILL BE PERFORMED WITH VAR. 4 AS THE DEPENDENT VARIABLE, AND VARIABLES 3, 2 AND 1 AS THE INDEPENDENT VARIABLES.

STEP-WISE MULTIPLE REGRESSION IS A METHOD OF DECIDING WHICH INDEPENDENT VARIABLES ARE IMPORTANT IN TERMS OF ACCOUNTING FOR VARIATION IN THE DEPENDENT VARIABLE Y. INDEPENDENT VARIABLES ARE ENTERED INTO THE REGRESSION SEQUENTIALLY, AND AS EACH VARIABLE IS ENTERED, A TEST MAY BE MADE TO DECIDE IF THE RESIDUAL SUM OF SQUARES HAS BEEN REDUCED SIGNIFICANTLY.

IN THE FIRST STEP, THE INDEPENDENT VARIABLE WHICH IS MOST HIGHLY CORRELATED WITH THE DEPENDENT VARIABLE IS ENTERED INTO THE REGRESSION EQUATION. OUTPUT INCLUDES SUM OF SQUARES REDUCED BY INCLUSION OF THIS INDEPENDENT VARIABLE AND THE PROPORTION OF THE TOTAL REDUCED. AN F-TEST IS PERFORMED ON THIS REGRESSION AND THE SIGNIFICANCE LEVEL OF THE CALCULATED F-STATISTIC IS PRINTED. A DECISION MAY THEN BE MADE AS TO WHETHER THE REGRESSION OF Y ON THE INDEPENDENT VARIABLE MOST HIGHLY CORRELATED WITH Y IS SIGNIFICANT. IF IT IS DECIDED THAT THIS REGRESSION IS SIGNIFICANT, PROCEED TO STEP 2.

IN THE SECOND STEP, THE INDEPENDENT VARIABLE WHICH HAS THE HIGHEST PARTIAL CORRELATION WITH Y (ALLOWING FOR THE INDEPENDENT VARIABLE ENTERED IN THE FIRST STEP) IS ENTERED INTO THE REGRESSION EQUATION. A DECISION MAY BE MADE WHETHER THE NEWLY ENTERED VARIABLE MAKES A SIGNIFICANT REDUCTION IN THE RESIDUAL SUM OF SQUARES BY REFERRING TO THE SIGNIFICANCE LEVEL OF THE COMPUTED T-VALUE FOR THE REGRESSSION COEFFICIENT OF THE NEWLY ENTERED VARIABLE. IF IT IS DECIDED THAT THE NEWLY ENTERED VARIABLE MAKES A SIGNIFICANT CONTRIBUTION TO REDUCING THE RESIDUAL SUM OF SQUARES, PROCEED TO STEP 3.

IN STEP 3 AND SUBSEQUENT STEPS, THE INDEPENDENT VARIABLE WITH THE HIGHEST PARTIAL CORRELATION WITH THE DEPENDENT VARIABLE (ALLOWING FOR PREVIOUSLY ENTERED INDEPENDENT VARIABLES) IS ENTERED INTO THE REGRESSION EQUATION. A DECISION MAY BE MADE WHETHER THE NEWLY ENTERED VARIABLE MAKES A SIGNIFICANT REDUCTION IN THE RESIDUAL SUM OF SQUARES OVER THE REDUCTION PRODUCED BY VARIABLES ALREADY IN THE EQUATION BY REFERRING TO THE SIGNIFICANCE LEVEL OF THE COMPUTED T-VALUE OF THE NEWLY ENTERED VARIABLE. WHEN IT IS DECIDED THAT A NEWLY ENTERED VARIABLE DOES NOT TAKE UP A SIGNIFICANT AMOUNT OF VARIATION IN Y, BEYOND THAT ACCOUNTED FOR BY PREVIOUSLY ENTERED VARIABLES, THE NEWLY ENTERED VARIABLE AND ALL VARIABLES NOT ALREADY IN THE REGRESSION EQUATION ARE NOT INCLUDED IN THE REGRESSION EQUATION.

THE REGRESSION COEFFICIENTS AND F-STATISTICS FOR THE REGRESSION EQUATION WILL BE THOSE IN THE STEP PRECEDING THE STEP IN WHICH IT IS DECIDED NOT TO INCLUDE A VARIABLE IN THE EQUATION.

REFERENCES

DRAPER, N.R. AND H. SMITH (1966) **APPLIED REGRESSION ANALYSIS.** JOHN **WILEY AND SONS,** NEW YORK. PP 169-171.

POLYNOMIAL REGRESSION

```
1      8
$POLRG I1,I2,I3
```

PARAMETER FIELD (COLUMNS 8-79)

I1 - DEPENDENT VARIABLE

I2 - INDEPENDENT VARIABLE

I3 - DEGREE POLYNOMIAL TO BE FITTED

OUTPUT

1. THE NUMBER OF CASES IS GIVEN.
2. THE INTERCEPT AND REGRESSION COEFFICIENTS ARE PRINTED.
3. AN ANALYSIS OF VARIANCE FOR THE REGRESSION IS COMPUTED.

USAGE

1. ONLY ONE DEPENDENT AND ONE INDEPENDENT VARIABLE MAY BE USED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.
3. A 5TH DEGREE POLYNOMIAL OR A LESSER DEGREE MAY BE FITTED, ($I3 \leq 5$).

EXAMPLE

CARDS	COMMENTS
1 8	
$POLRG 2,4,3	A 3RD DEGREE POLYNOMIAL WILL BE FITTED USING VAR. 2 AS THE DEPENDENT VARIABLE AND VAR. 4 AS THE INDEPENDENT VARIABLE.

IN THIS ROUTINE, THE METHOD OF LEAST SQUARES IS USED TO ESTIMATE THE BJ↑S, J=0,...,5, IN ANY OF THE FOLLOWING FIRST THROUGH FIFTH DEGREE POLYNOMIAL MODELS?

```
Y(I)=B0 + B1 X(I) + E(I)
Y(I)=B0 + B1 X(I) + B2 [X(I)↑2] + E(I)
Y(I)=B0 + B1 X(I) + B2 [X(I)↑2] + B3 [X(I)↑3] + E(I)
Y(I)=B0 + B1 X(I) + B2 [X(I)↑2] + B3 [X(I)↑3] + B4 [X(I)↑4] + E(I)
Y(I)=B0 + B1 X(I) + B2 [X(I)↑2] + B3 [X(I)↑3] + B4 [X(I)↑4] +
     B5 [X(I)↑5] + E(I)
```

WHERE

X(I) AND Y(I) ARE OBSERVATIONS OF THE INDEPENDENT VARIABLE X AND THE
DEPENDENT VARIABLE Y RESPECTIVELY

BJ↑S ARE UNKNOWN PARAMETERS TO BE ESTIMATED AND

E(I) IS A RANDOM ERROR.

IT IS ASSUMED THAT THE E(I)↑S ARE INDEPENDENT AND NORMALLY DISTRIBUTED WITH MEANS ZERO AND COMMON VARIANCE.

ONLY ONE MODEL MAY BE CHOSEN WITH ANY **$POLRG** CARD. IT WILL BE NOTED THAT IF A POLYNOMIAL OF DEGREE 1 IS CHOSEN, **$POLRG** WILL PERFORM THE SAME CALCULATION AS **$LSQR**.

REFERENCE

DRAPER, N.R. AND H. SMITH (1966) **APPLIED REGRESSION ANALYSIS.** JOHN WILEY AND SONS, NEW YORK. PP 129-130.

NORMALITY TEST

```
1      8
$NORMT I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
    I1
    I2
    ..
    ..    INPUT VARIABLES
    ..
    I50
```

OUTPUT

FOR EACH VARIABLE SPECIFIED THE OUTPUT INCLUDES

1. SUM, MEAN, STANDARD DEVIATION, STANDARD ERROR OF THE MEAN, NUMBER OF OBSERVATIONS, MINIMUM, MAXIMUM, AND RANGE;
2. THE SECOND MOMENT ABOUT THE MEAN, AND THE 3RD AND 4TH MOMENTS RELATIVE TO THE 2ND MOMENT;
3. COEFFICIENT OF VARIATION;
4. GEARY'S KURTOSIS TEST STATISTIC;
5. HISTOGRAM WITH OBSERVED AND EXPECTED FREQUENCIES, AND
6. CHI-SQUARE TEST FOR GOODNESS OF FIT TO THE NORMAL DISTRIBUTION.

USAGE

1. A MAXIMUM OF 50 VARIABLES MAY BE USED.
2. THERE IS NO LIMIT ON THE NUMBER OF OBSERVATIONS.

EXAMPLE

CARDS	COMMENTS
1 8	
$NORMT 1,6	NORMALITY TESTS WILL BE PERFORMED ON VARIABLES 1 AND 6.

THIS PROGRAM CONTAINS FOUR TEST STATISTICS WHICH MAY BE USED TO DECIDE IF THE OBSERVATIONS OF INTEREST ARE NORMALLY DISTRIBUTED.

1. THE RELATIVE THIRD MOMENT ABOUT THE MEAN,

$$G(1) = \frac{SUM[(X(I)-XBAR)\uparrow 3]/N}{[(SUM[(X(I)-XBAR)\uparrow 2]/N)\ \uparrow(3/2)]},$$

IS A MEASURE OF THE SKEWNESS OF THE DISTRIBUTION OF THE OBSERVATIONS. FOR A NORMAL DISTRIBUTION G(1) = 0. TABLE A.6 IN SNEDECOR AND COCHRAN (1967) MAY BE USED FOR DECIDING IF G(1) DIFFERS SIGNIFICANTLY FROM 0.

2. THE RELATIVE FOURTH MOMENT ABOUT THE MEAN,

$$G(2) = \frac{SUM[(X(I)-XBAR)\uparrow 4]/N}{[SUM[(X(I)-XBAR)\uparrow 2]/N]\uparrow 2},$$

IS A MEASURE OF THE KURTOSIS OF THE DISTRIBUTION OF THE OBSERVATION. FOR A NORMAL DISTRIBUTION G(2) = 3. WHEN G(2) IS LESS THAN 3, THE DISTRIBUTION IS SAID TO BE PLATYKURTIC, FLATTER THAN THE NORMAL CURVE. WHEN G(2) IS GREATER THAN 3, THE DISTRIBUTION IS SAID TO BE LEPTOKURTIC, MORE PEAKED THAN THE NORMAL CURVE. TABLE A.6 IN SNEDECOR AND COCHRAN (1967) MAY BE USED FOR DECIDING IF G(2) DIFFERS SIGNIFICANTLY FROM 3.

3. AN ALTERNATE METHOD FOR TESTING KURTOSIS IS **GEARY↑S TEST STATISTIC,**

$$A = \frac{[SUM[ABS(X(I)-XBAR)]]/N}{SQRT[(SUM[(X(I)-XBAR)\uparrow 2])/N]},$$

FOR NORMAL DISTRIBUTIONS, A = .7979. TO DECIDE IF THE OBTAINED **GEARY↑S KURTOSIS STATISTIC** DIFFERS SIGNIFICANTLY FROM .7979, SEE GEARY (1936).

4. A CHI-SQUARE GOODNESS OF FIT STATISTIC IS USED TO DECIDE IF THE OBSERVED DISTRIBUTION FITS A NORMAL DISTRIBUTION. THE GENERAL FORMULA IS

```
             SUM((OBSERVED - EXPECTED)↑2)
CHI-SQUARE = ---------------------------- ,
                       EXPECTED
```

WITH K-3 DEGREES OF FREEDOM. K IS THE NUMBER OF CLASSES USED IN CALCULATING CHI-SQUARE, OBSERVED IS THE OBSERVED FREQUENCY OF EACH CLASS, AND EXPECTED IS THE EXPECTED FREQUENCY OF EACH CLASS IF THE DISTRIBUTION OF THE OBSERVATIONS IS NORMAL.

REFERENCES

1. SNEDECOR, G.W. AND W.G. COCHRAN (1967) **STATISTICAL METHODS.** IOWA STATE UNIV. PRESS, AMES, IOWA. PP 84-90.

2. GEARY, R.C. (1936) **BIOMETRIKA,** VOL 28? P 295.

HISTOGRAM

```
1      8
$HSTGM I1
```

PARAMETER FIELD (COLUMNS 8-79)

I1 - INPUT VARIABLE

OUTPUT

1. THE MEAN, STANDARD DEVIATION, NUMBER OF OBSERVATIONS, MINIMUM, MAXIMUM AND RANGE BEFORE AND AFTER SCALING AND THE MEDIAN BEFORE SCALING ARE PRINTED.
2. THE LENGTHS AND MID-POINTS OF THE INTERVALS ARE GIVEN.
3. THE HISTOGRAM IS PLOTTED USING 20 INTERVALS.

USAGE

1. ONLY ONE INPUT VARIABLE MAY BE USED.
2. THERE IS A LIMIT OF 1000 CASES.
3. AUTO-SCALING WILL BE USED IF THE $SCALE CARD DOES NOT PRECEDE THE $HSTGM CARD.
4. REFER TO THE $SCALE DOCUMENTATION FOR MANUALLY SCALING HISTOGRAMS.

EXAMPLE

CARD	COMMENT
1 8	
$HSTGM 3	A HISTOGRAM OF VARIABLE 3 WILL BE PLOTTED.

PLOT ROUTINE

```
1      8
$PLOT  I1,I2
```

PARAMETER FIELD (COLUMNS 8-79)

I1 - DEPENDENT VARIABLE (VERTICAL AXIS)

I2 - INDEPENDENT VARIABLE (HORIZONTAL AXIS)

OUTPUT

A PLOT OF VARIABLE I1 VERSUS VARIABLE I2 IS PRINTED.

USAGE

1. ONLY ONE DEPENDENT AND ONE INDEPENDENT VARIABLE MAY BE USED.
3. OVERLAYING POINTS ARE PLOTTED WITH A $ SYMBOL.
4. AUTO-SCALING IS USED IF THE $SCALE CARD DOES NOT PRECEDE THE $PLOT CARD.
5. REFER TO THE $SCALE DOCUMENTATION FOR MANUAL SCALING OF THE AXES.

EXAMPLE

CARD		COMMENT
1	8	
$PLOT	2,4	DEPENDENT VARIABLE 2 WILL BE PLOTTED AGAINST INDEPENDENT VARIABLE 4.

PLOT NORMAL

```
1      8
$PLOTN I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
     I1 - DEPENDENT VARIABLE (VERTICAL AXIS)
     I2 -
     ..
     ..   INDEPENDENT VARIABLES (HORIZONTAL AXIS)
     ..
     I50 -
```

OUTPUT

A PLOT IS MADE OF THE DEPENDENT VARIABLE VERSUS EACH OF THE INDEPENDENT VARIABLES USING THE SYMBOL -A- FOR I1 VERSUS I2, THE SYMBOL -B- FOR I1 VERSUS I3, ETC., THROUGH THE CDC FORTRAN CHARACTER SET.

USAGE

1. A MAXIMUM OF 1 DEPENDENT AND 49 INDEPENDENT VARIABLES MAY BE USED.
2. NO MORE THAN 2500 DATA POINTS (INCLUDING MISSING DATA) ARE ALLOWED.
3. AUTO-SCALING IS USED IF THE $SCALE CARD DOES NOT PRECEDE THE $PLOTN CARD.
4. REFER TO THE $SCALE DOCUMENTATION FOR MANUAL SCALING OF THE AXES.

EXAMPLES

CARDS	COMMENTS
1 8	
$PLOTN 4,1,2	PLOT OF DEPENDENT VARIABLE 4 VERSUS INDEPENDENT VARIABLES 1 AND 2
$PLOTN ALL	PLOT OF THE 1ST VARIABLE (DEPENDENT) VERSUS THE REMAINING VARIABLES AS INDEPENDENT

PLOT INVERTED

```
1      8
$PLOTI I1,I2,...,I50
```

PARAMETER FIELD (COLUMNS 8-79)

```
I1 - INDEPENDENT VARIABLE (HORIZONTAL AXIS)
I2 -
..
..   DEPENDENT VARIABLES (VERTICAL AXIS)
..
I50 -
```

OUTPUT

A PLOT IS MADE OF THE INDEPENDENT VARIABLE VERSUS EACH OF THE DEPENDENT VARIABLES USING THE SYMBOL -A- FOR I1 VERSUS I2, THE SYMBOL -B- FOR I1 VERSUS I3, ETC., THROUGH THE CDC FORTRAN CHARACTER SET.

USAGE

1. A MAXIMUM OF 1 INDEPENDENT VARIABLE AND 49 DEPENDENT VARIABLES MAY BE USED.
2. NO MORE THAN 2500 DATA POINTS (INCLUDING MISSING DATA) ARE ALLOWED.
3. AUTO-SCALING IS USED IF THE $SCALE CARD DOES NOT PRECEDE THE $PLOTI CARD.
4. REFER TO THE $SCALE DOCUMENTATION FOR MANUAL SCALING OF THE AXES.
5. THE PLOT SHOULD BE ROTATED 90 DEGREES COUNTERCLOCKWISE FOR PROPER ALIGNMENT OF THE AXES.

EXAMPLES

CARDS	COMMENTS
1 8	
$PLOTI 7,1,4	PLOT OF INDEPENDENT VARIABLE 7 VERSUS DEPENDENT VARIABLES 1 AND 4
$PLOTI ALL	PLOT OF THE 1ST VARIABLE (INDEPENDENT) VERSUS THE REMAINIG VARIABLES AS DEPENDENT

F-DISTRIBUTION

```
1     8
$FDIST I1,I2,I3
```

PARAMETER FIELD (COLUMNS 8-79)

I1 - VARIABLE CONTAINING THE NUMERATOR DEGREES OF FREEDOM - R1

I2 - VARIABLE CONTAINING THE DENOMINATOR DEGREES OF FREEDOM - R2

I3 - VARIABLE CONTAINING THE F-STATISTIC

OUTPUT

THE PROBABILITY OF FINDING A RANDOM VALUE OF -F-, WITH R1 AND R2 DEGREES OF FREEDOM, GREATER THAN THE INPUT F-VALUE IS PRINTED AS THE LEVEL SIGNIFICANCE.

EXAMPLE

CARD	COMMENTS
1 8	
$FDIST 3,7,9	R1 IS IN VARIABLE 3.
	R2 IS IN VARIABLE 7.
	THE F-STATISTIC IS IN VARIABLE 9.

CHI-SQUARE DISTRIBUTION

```
1      8
$CDIST I1,I2
```

PARAMETER FIELD (CCLUMNS 8-79)

I1 - VARIABLE CONTAINING THE DEGREES OF FREEDOM

I2 - VARIABLE CONTAINING THE CHI-SQUARE STATISTIC

OUTPUT

THE PROBABILITY OF FINDING A RANDOM VALUE OF CHI-SQUARE, WITH THE GIVEN DEGREES OF FREEDOM, GREATER THAN THE INPUT CHI-SQUARE IS PRINTED AS THE LEVEL OF SIGNIFICANCE.

EXAMPLE

CARD	COMMENT
1 8	
$CDIST 4,6	THE DEGREES OF FREEDOM ARE IN VARIABLE 4. THE CHI-SQUARE STATISTIC IS IN VARIABLE 6.

T-DISTRIBUTION

```
1      8
$TDIST I1,I2
```

PARAMETER FIELD (COLUMNS 8-79)

I1 - VARIABLE CONTAINING THE DEGREES OF FREEDOM

I2 - VARIABLE CONTAINING THE T-STATISTIC

OUTPUT

TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF -T-, WITH THE GIVEN DEGREES OF FREEDOM, GREATER THAN THE ABSOLUTE VALUE OF THE INPUT T VALUE IS PRINTED AS THE LEVEL OF SIGNIFICANCE.

EXAMPLE

CARD	COMMENT
1 8	
$TDIST 1,2	DEGREES OF FREEDOM ARE IN VARIABLE 1. T-VALUE IS IN VARIABLE 2.

WILCOXON'S RANK SUM AND SIGNED RANKS DISTRIBUTIONS

```
1      8
$RSDST I1,I2,I3,
```

PARAMETER FIELD (COLUMNS 8-79)

- I1 - VARIABLE CONTAINING THE NUMBER OF OBSERVATIONS -M- OF THE FIRST VARIABLE (X)
- I2 - VARIABLE CONTAINING THE NUMBER OF OBSERVATIONS -N- OF THE SECOND VARIABLE (Y)
- I3 - VARIABLE CONTAINING WILCOXON'S RANK SUM OR SIGNED RANKS STATISTIC

OUTPUT

LEVEL OF SIGNIFICANCE OF THE INPUT WILCOXON'S RANK SUM OR SIGNED RANKS STATISTIC IS PRINTED.

USAGE

1. WHEN INPUTTING THE WILCOXON'S SIGNED RANKS STATISTIC, SET M=0 AND N SHOULD BE THE NUMBER OF DIFFERENCES.
2. THIS ROUTINE IS LIMITED TO M+N≤30.

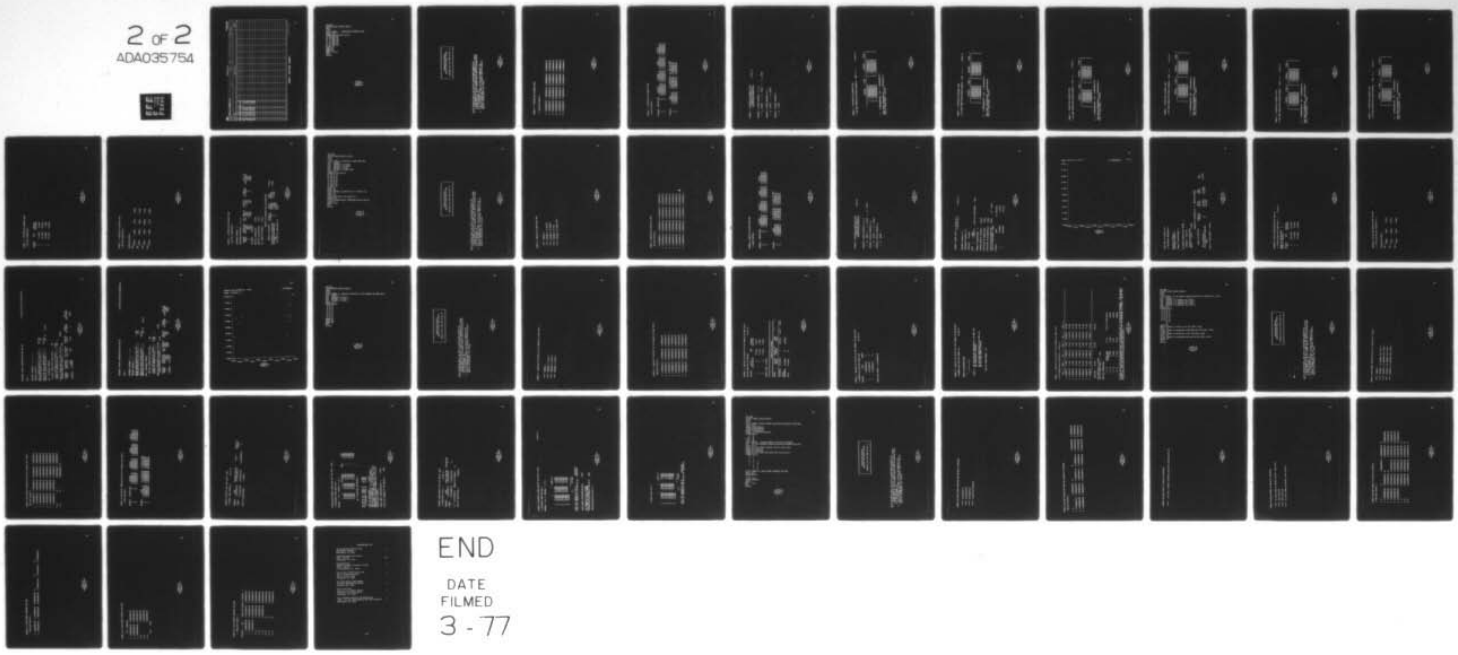
EXAMPLE

CARD	COMMENT
1 8	
$RSDST 3,4,6	M IS IN VARIABLE 3. N IS IN VARIABLE 4. THE WILCOXON STATISTIC IS IN VARIABLE 6.

AD-A035 754 LETTERMAN ARMY INST OF RESEARCH SAN FRANCISCO CALIF F/G 12/1
GENERALIZED RESEARCH ANALYSIS STATISTICAL SYSTEM. SECOND EDITIO--ETC(U)
NOV 76 H LAZARUS, T MCCAA, R TEPLICK

UNCLASSIFIED LAIR-33 NL

2 OF 2
ADA035754



END
DATE
FILMED
3-77

IBM

FORTRAN Coding Form

GX28-7327-6 U/M 050**
Printed in U.S.A.

PROGRAM GRASS-EXAMPLE #1

PROGRAMMER | DATE | PUNCHING INSTRUCTIONS | GRAPHIC | PUNCH | PAGE OF | CARD ELECTRO NUMBER*

FORTRAN STATEMENT

```
AGE TIME TEMP PULSE

20 10.0 999  123
18 15.8 997  147
35  8.8 991  109
24  8.0 986  118
41 12.5 994  131
43 12.8 997  135
```

*A standard card form, IBM electro 888157, is available for punching statements from this form

**Number of forms per pad may vary slightly

FIGURE 1 DATA SHEET EXAMPLE 1

```
JOB CARD
FETCHPS (GRASZ,GRAZZ,GRASS)
GRAZZ.
7/8/9
$TITLE EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY
$DATA  4,6,1
$FORMAT(F2.0,F5.0,F4.1,F4.0)
20 10.0 989 123
18 15.8 987 147
35  8.8 991 109
24  8.0 986 118
41 12.5 994 131
43 12.8 997 135
$BSTAT ALL
$CORR ALL
$RNKCR ALL
$MREG  2,1,3,4
$END
6/7/8/9
```

FIGURE 2
DECK SETUP
EXAMPLE 1

```
=======================================
*                                     *
*        GENERALIZED RESEARCH         *
*    ANALYSIS STATISTICAL SYSTEM      *
*                                     *
*  VERSION 4.00, LAST MOD 05 DEC 74   *
*                                     *
=======================================
```

UNLESS OTHERWISE NOTED, THE LEVEL OF SIGNIFICANCE PRINTED FOR

1. F OR CHI-SQUARE STATISTICS IS EQUAL TO THE PROBABILITY OF FINDING A RANDOM VALUE OF F OR CHI-SQUARE GREATER THAN THE OBTAINED VALUE OF F OR CHI-SQUARE, OR
2. T OR Z STATISTICS IS EQUAL TO TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T OR Z GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T OR Z VALUE.

FIGURE 3
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---RAW DATA DISPLAY---

1	2.00000000E+01	1.00000000E+01	9.89000000E+01	1.23000000E+02
2	1.80000000E+01	1.58000000E+01	9.87000000E+01	1.47000000E+02
3	3.50000000E+01	8.80000000E+00	9.91000000E+01	1.09000000E+02
4	2.40000000E+01	8.00000000E+00	9.86000000E+01	1.18000000E+02
5	4.10000000E+01	1.25000000E+01	9.94000000E+01	1.31000000E+02
6	4.30000000E+01	1.28000000E+01	9.97000000E+01	1.35000000E+02

FIGURE 4
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---BSTAT ROUTINE---

VARIABLE	OBSERVATIONS	MINIMUM	MAXIMUM	RANGE	MEDIAN
1	6	1.80000000E+01	4.30000000E+01	3.80000000E+01	2.95000000E+01
2	6	8.00000000E+00	1.58000000E+01	3.80000000E+01	1.12500000E+01
3	6	9.86000000E+01	9.97000000E+01	3.80000000E+01	9.90000000E+01
4	6	1.09000000E+02	1.47000000E+02	3.80000000E+01	1.27000000E+02

VARIABLE	MEAN	ST. DEVIATION	ST. ERR. OF MEAN
1	3.01666667E+01	1.09071842E+01	4.45283929E+00
2	1.13166667E+01	2.92466522E+00	1.19398958E+00
3	9.90666667E+01	4.22689800E-01	1.72562388E-01
4	1.27166667E+02	1.34226177E+01	5.47976074E+00

FIGURE 5
COMPUTER OUTPUT
EXAMPLE 1

```
EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY
            ----CORRELATION ROUTINE----          05 DEC 74
             (OBSERVATIONS,CORRELATION)

VARIABLE  1 VERSUS VARIABLES  2, 3, 4,

   (   6, -.00637)   (   6,  .90810)   (   6, -.12044)

VARIABLE  2 VERSUS VARIABLES 3, 4,

   (   6,  .22218)   (   6, .95109)

VARIABLE  3 VERSUS VARIABLES 4,

   (   6,  .11750)
```

FIGURE 6
COMPUTER OUTPUT
EXAMPLE 1

```
EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---THE RANK CORRELATION BETWEEN VARIABLES  1 AND  2 ---            05 DEC 74


               VARIABLE     1 RANK  1        VARIABLE     2 RANK  2
   -------------------------------------------------------------------
       1  2.0000000000E+01           2  1.0000000000E+01          3
       2  1.8000000000E+01           1  1.5800000000E+01          6
       3  3.5000000000E+01           4  8.8000000000E+00          2
       4  2.4000000000E+01           3  8.0000000000E+00          1
       5  4.1000000000E+01           5  1.2500000000E+01          4
       6  4.3000000000E+01           6  1.2800000000E+01          5

S =   1.0000000000E+00;  TAU =    6.6666666667E-02

LARGE SAMPLE APPROXIMATION =    1.8786728733E-01
LEVEL OF SIGNIFICANCE = .8509808
```

FIGURE 7
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---THE RANK CORRELATION BETWEEN VARIABLES 1 AND 3 ---- 05 DEC 74

	VARIABLE 1	RANK 1	VARIABLE 3	RANK 3
1	2.0000000000E+01	2	9.8900000000E+01	3
2	1.8000000000E+01	1	9.8700000000E+01	2
3	3.5000000000E+01	4	9.9100000000E+01	4
4	2.4000000000E+01	3	9.8600000000E+01	1
5	4.1000000000E+01	5	9.9400000000E+01	5
6	4.3000000000E+01	6	9.9700000000E+01	6

S = 1.1000000000E+01; TAU= 7.3333333333E-01

LARGE SAMPLE APPROXIMATION= 2.0665401606E+00
LEVEL OF SIGNIFICANCE = .0387778

FIGURE 8
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---THE RANK CORRELATION BETWEEN VARIABLES 1 AND 4 --- 05 DEC 74

	VARIABLE 1	RANK 1	VARIABLE 4	RANK 4
1	2.0000000000E+01	2	1.2300000000E+02	3
2	1.8000000000E+01	1	1.4700000000E+02	6
3	3.5000000000E+01	4	1.0900000000E+02	1
4	2.4000000000E+01	3	1.1800000000E+02	2
5	4.1000000000E+01	5	1.3100000000E+02	4
6	4.3000000000E+01	6	1.3500000000E+02	5

S = -1.0000000000E+00; TAU= -6.6666666667E-02

LARGE SAMPLE APPROXIMATION = -1.8786728733E-01
LEVEL OF SIGNIFICANCE = .8509808

FIGURE 9
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---THE RANK CORRELATION BETWEEN VARIABLES 2 AND 3 --- 05 DEC 74

	VARIABLE 2	RANK 2	VARIABLE 3	RANK 3
1	1.0000000000E+01	3	9.8900000000E+01	3
2	1.5800000000E+01	6	9.8700000000E+01	2
3	8.8000000000E+00	2	9.9100000000E+01	4
4	8.0000000000E+00	1	9.8600000000E+01	1
5	1.2500000000E+01	4	9.9400000000E+01	5
6	1.2800000000E+01	5	9.9700000000E+01	6

S = 5.0000000000E+00; TAU= 3.3333333333E-01

LARGE SAMPLE APPROXIMATION = 9.3933643663E-01
LEVEL OF SIGNIFICANCE = .3475582

FIGURE 10
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---THE RANK CORRELATION BETWEEN VARIABLES 2 AND 4 --- 05 DEC 74

	VARIABLE 2	RANK 2	VARIABLE 4	RANK 4
1	1.0000000000E+01	3	1.2300000000E+02	3
2	1.5800000000E+01	6	1.4700000000E+02	6
3	8.8000000000E+00	2	1.0900000000E+02	1
4	8.0000000000E+00	1	1.1800000000E+02	2
5	1.2500000000E+01	4	1.3100000000E+02	4
6	1.2800000000E+01	5	1.3500000000E+02	5

S = 1.3000000000E+01; TAU= 8.6666666667E-01

LARGE SAMPLE APPROXIMATION= 2.4422747352E+00
LEVEL OF SIGNIFICANCE= .0145950

FIGURE 11
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---THE RANK CORRELATION BETWEEN VARIABLES 3 AND 4 --- 05 DEC 74

	VARIABLE 3	RANK 3	VARIABLE 4	RANK 4
1	9.8900000000E+01	3	1.2300000000E+02	3
2	9.8700000000E+01	2	1.4700000000E+02	6
3	9.9100000000E+01	4	1.0900000000E+02	1
4	9.8600000000E+01	1	1.1800000000E+02	2
5	9.9400000000E+01	5	1.3100000000E+02	4
6	9.9700000000E+01	6	1.3500000000E+02	5

S = 3.0000000000E+00; TAU= 2.0000000000E-01

LARGE SAMPLE APPROXIMATION= 5.6360186198E-01
LEVEL OF SIGNIFICANCE= .5730251

FIGURE 12
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---MULTIPLE REGRESSION---

VARIABLE NO.	MEAN	STANDARD DEVIATION
2	11.31667	2.92467
1	30.16667	10.90718
3	99.06667	.42269
4	127.16667	13.42262

FIGURE 13
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---MULTIPLE REGRESSION---

CORRELATION MATRIX

VAR 2	1.00000	-.00637	.22218	.95109
VAR 1	-.00637	1.00000	.90810	-.12044
VAR 3	.22218	.90810	1.00000	.11750
VAR 4	.95109	-.12044	.11750	1.00000

FIGURE 14
COMPUTER OUTPUT
EXAMPLE 1

EXAMPLE 1 - HYPOTHETICAL RESEARCH STUDY

---MULTIPLE REGRESSION---

DEPENDENT VARIABLE 2

INDEPENDENT VARIABLES 1, 3, 4,

VARIABLE NUMBER	REGRESSION COEFFICIENT	STD. ERROR OF REG. COEFF.	COMPUTED T- VALUE	DEGREES OF FREEDOM	LEVEL OF SIGNIFICANCE
4	.20637	.05332	3.871	2	.06073
3	.51636	4.01381	.129	2	.90941
1	.01071	.15560	.069	2	.95140
INTERCEPT	-66.40421				

MULTIPLE CORRELATION .95767

MULTIPLE DETERMINATION .91714

STD ERROR OF ESTIMATE 1.33116

ANALYSIS OF VARIANCE FOR THE REGRESSION

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F VALUE
ATTRIBUTABLE TO REGRESSION	3	39.22434	13.07478	7.37856
DEVIATION FROM REGRESSION	2	3.54399	1.77200	
TOTAL	5	42.76833		

THE LEVEL OF SIGNIFICANCE IS .02766

FIGURE 15
COMPUTER OUTPUT
EXAMPLE 1

```
JOB CARD
FETCHPS( GRASZ,GRAZZ,GRASS)
GRAZZ.
7/8/9
$TITLE EXAMPLE 2, ANALYSIS OF HEART RATE DATA
$CMT   EXAMPLE2
$CMT   VARIABLE 1 IS HEIGHT
$CMT   VARIABLE 2 IS WEIGHT
$CMT   VARIABLE 3 IS AGE
$CMT   VARIABLE 4 IS HEART RATE
$DATA  4,8,1
$FORMAT(F2.0,1X,3F3.0)
68 150 22 70
69 160 23 78
70 169 25 79
71 167 27 82
72 170 29 83
73 175 24 87
74 180 23 89
75 180 21 72
$BSTAT ALL
$CORR  ALL
$TITLE (EXAMPLE 2) HEART RATE IS 4, WEIGHT IS 2
$LSQR  4,2
$PLOT  4,2
$TITLE HEART RATE IS 4, HEIGHT IS 1
$POLRG 4,1,3
$TITLE DEPENDENT-WEIGHT INDEPENDENT-HEIGHT AND AGE
$SWREG 2,1,3
$PLOTN 2,1,3
$END
6/7/8/9
```

FIGURE 16
DECK SETUP
EXAMPLE 2

```
========================================
*                                      *
*        GENERALIZED RESEARCH          *
*    ANALYSIS STATISTICAL SYSTEM       *
*                                      *
*   VERSION 4.00, LAST MOD 05 DEC 74   *
*                                      *
========================================
```

UNLESS OTHERWISE NOTED, THE LEVEL OF SIGNIFICANCE PRINTED FOR

1. F OR CHI-SQUARE STATISTICS IS EQUAL TO THE PROBABILITY OF FINDING A RANDOM VALUE OF F OR CHI-SQUARE GREATER THAN THE OBTAINED VALUE OF F OR CHI-SQUARE, OR
2. T OR Z STATISTICS IS EQUAL TO TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T OR Z GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T OR Z VALUE.

FIGURE 17
COMPUTER OUTPUT
EXAMPLE 2

EXAMPLE 2, ANALYSIS OF HEART RATE DATA

```
*****  EXAMPLE 2
*****  VARIABLE 1 IS HEIGHT
*****  VARIABLE 2 IS WEIGHT
*****  VARIABLE 3 IS AGE
*****  VARIABLE 4 IS HEART RATE
```

FIGURE 18
COMPUTER OUTPUT
EXAMPLE 2

EXAMPLE 2, ANALYSIS OF HEART RATE DATA

RAW DATA DISPLAY---

1	6.80000000E+01	1.50000000E+02	2.20000000E+01	7.00000000E+01
2	6.90000000E+01	1.60000000E+02	2.30000000E+01	7.80000000E+01
3	7.00000000E+01	1.69000000E+02	2.50000000E+01	7.90000000E+01
4	7.10000000E+01	1.67000000E+02	2.70000000E+01	8.20000000E+01
5	7.20000000E+01	1.70000000E+02	2.90000000E+01	8.30000000E+01
6	7.30000000E+01	1.75000000E+02	2.40000000E+01	8.70000000E+01
7	7.40000000E+01	1.80000000E+02	2.30000000E+01	8.90000000E+01
8	7.50000000E+01	1.80000000E+02	2.10000000E+01	7.20000000E+01

FIGURE 19
COMPUTER OUTPUT
EXAMPLE 2

EXAMPLE 2, ANALYSIS OF HEART RATE DATA

---BSTAT ROUTINE---

VARIABLE	OBSERVATIONS	MINIMUM	MAXIMUM	RANGE	MEDIAN
1	8	6.80000000E+01	7.50000000E+01	1.90000000E+01	7.15000000E+01
2	8	1.50000000E+02	1.80000000E+02	1.90000000E+01	1.69500000E+02
3	8	2.10000000E+01	2.90000000E+01	1.90000000E+01	2.35000000E+01
4	8	7.00000000E+01	8.90000000E+01	1.90000000E+01	8.05000000E+01

VARIABLE	MEAN	ST. DEVIATION	ST. ERR. OF MEAN
1	7.15000000E+01	2.44948974E+00	8.66025404E-01
2	1.68875000E+02	1.01761275E+01	3.59780439E+00
3	2.42500000E+01	2.65921578E+00	9.40174756E-01
4	8.00000000E+01	6.67618368E+00	2.36038738E+00

FIGURE 20
COMPUTER OUTPUT
EXAMPLE 2

```
EXAMPLE 2, ANALYSIS OF HEART RATE DATA
             ---CORRELATION ROUTINE---                 13 DEC 74
             (OBSERVATIONS,CORRELATION)

VARIABLE 1 VERSUS VARIABLES 2, 3, 4,

   (   8,  .94851)   (   8, -.08773)   (   8,  .41058)
VARIABLE 2 VERSUS VARIABLES 3, 4,
   (   8,  .01188)   (   8,  .54672)
VARIABLE 3 VERSUS VARIABLES 4,
   (   8,  .45866)
```

FIGURE 21
COMPUTER OUTPUT
EXAMPLE 2

```
(EXAMPLE 2)HEART RATE IS 4, WEIGHT IS 2
                        ---LSQR ROUTINE---                    05 DEC 74

DEPENDENT VARIABLE IS 4

INDEPENDENT VARIABLE IS 2

INTERCEPT=     19.42749

SLOPE    =       .35868
T        =      1.59938 DF=    6 LEVEL OF SIGNIFICANCE=   .16085

NUMBER OF OBSERVATIONS=   8

STANDARD ERROR OF ESTIMATE=      6.03797

COEFFICIENT OF CORRELATION=       .54672

COEFFICIENT OF DETERMINATION=     .29890

F-RATIO FOR LSQR REGRESSION=     2.55801       DF = 1 AND 6
LEVEL OF SIGNIFICANCE=  .16085

VARIABLE             MEAN                 ST. DEVIATION

   4               80.00000                     6.67618
   2              168.87500                    10.17613
```

FIGURE 22
COMPUTER OUTPUT
EXAMPLE 2

(EXAMPLE 2)HEART RATE IS 4, WEIGHT IS 2 ---PLOT ROUTINE--- 05 DEC 74
VAR. 4 VERSUS VAR. 2

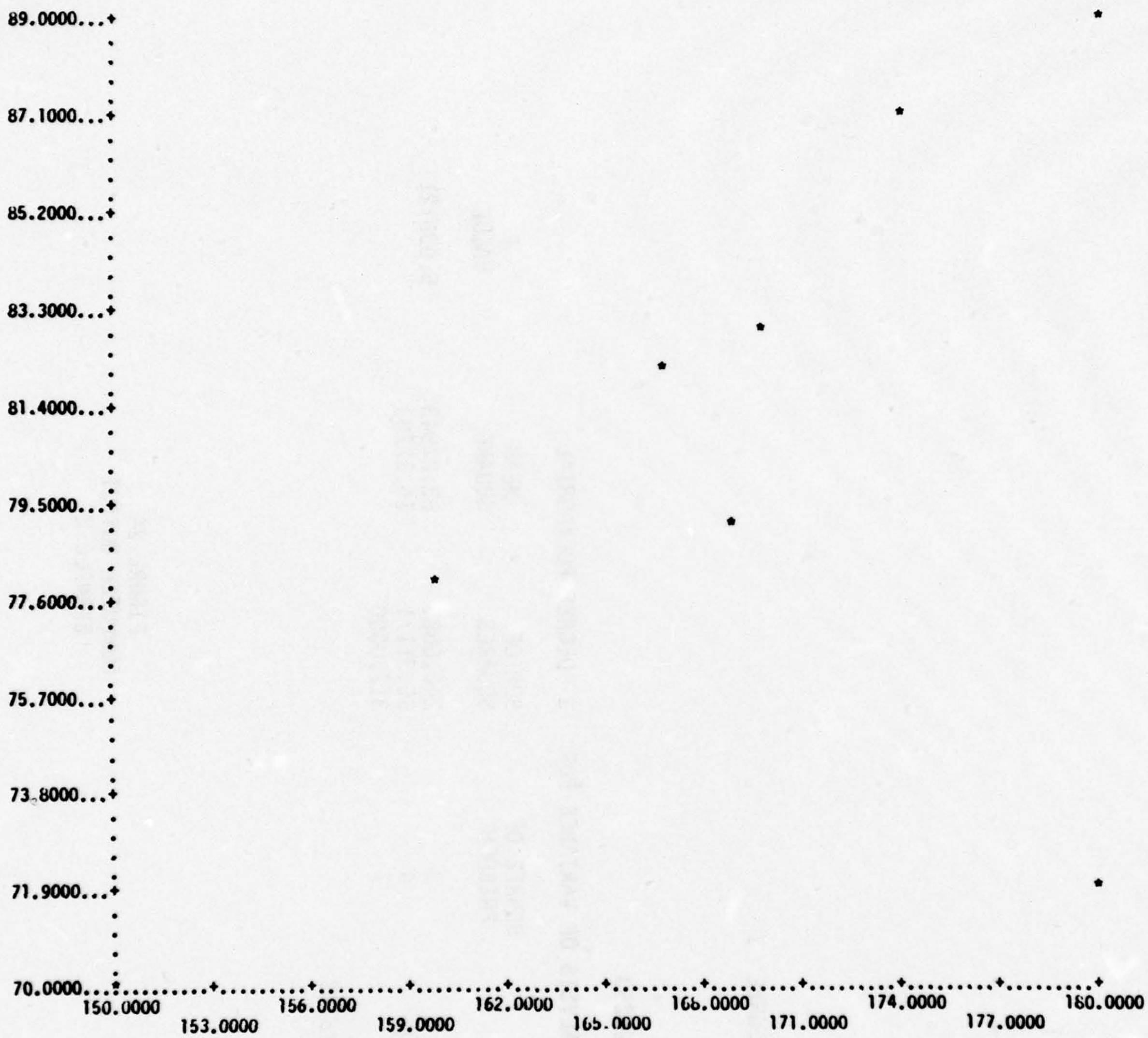


FIGURE 23
COMPUTER OUTPUT
EXAMPLE 2

HEART RATE IS 4,HEIGHT IS 1

---POLYNOMIAL REGRESSION---

DEPENDENT VARIABLE 4
INDEPENDENT VARIABLE 1

NUMBER OF OBSERVATIONS 8

POLYNOMIAL REGRESSION OF DEGREE 3

INTERCEPT 87212.879

REGRESSION COEFFICIENTS
-3731.6 53.202 -.25253

ANALYSIS OF VARIANCE FOR 3 DEGREE POLYNOMIAL

SOURCE OF VARIATION	DEGREE OF FREEDOM	SUM OF SQUARES	MEAN SQUARE	F VALUE
DUE TO REGRESSION	3	246.6883	82.22943	5.036121
DEVIATION ABOUT REGRESSION	4	65.31171	16.32793	
TOTAL	7	312.0000		

LEVEL OF SIGNIFICANCE = .076183

FIGURE 24
COMPUTER OUTPUT
EXAMPLE 2

DEPENDENT -WEIGHT INDEPENDENT-HEIGHT AND AGE

--- STEP-WISE MULTIPLE REGRESSION --- 05 DEC 74

NUMBER OF OBSERVATIONS 8
NUMBER OF VARIABLES 3

VARIABLE NO.	MEAN	STANDARD DEVIATION
2	168.87500	10.17613
1	71.50000	2.44949
3	24.25000	2.65922

FIGURE 25
COMPUTER OUTPUT
EXAMPLE 2

```
DEPENDENT -WEIGHT INDEPENDENT-HEIGHT AND AGE

--- STEP-WISE MULTIPLE REGRESSION ---

CORRELATION MATRIX

VAR  2
    1.00000      .94851      .01188

VAR  1
     .94851     1.00000     -.08773

VAR  3
     .01188     -.08773     1.00000
```

FIGURE 26
COMPUTER OUTPUT
EXAMPLE 2

```
DEPENDENT -WEIGHT INDEPENDENT-HEIGHT AND AGE                    ---STEP-WISE MULTIPLE REGRESSION---

STEP 1

DEPENDENT VARIABLE............ 2

VARIABLE ENTERED..... 1

SUM OF SQUARES REDUCED IN THIS STEP....      652.149
PROPORTION REDUCED IN THIS STEP........         .900

CUMULATIVE SUM OF SQUARES REDUCED......      652.149
CUMULATIVE PROPORTION REDUCED..........         .900  OF     724.875

FOR  1 VARIABLES ENTERED

  MULTIPLE CORRELATION COEFFICIENT....    .949

  F-VALUE FOR ANALYSIS OF VARIANCE.... 53.803
  DEGREES OF FREEDOM FOR AOV..........  1 AND   6

LEVEL OF SIGNIFICANCE OF F-VALUE FOR AOV               .00033
  STANDARD ERROR OF ESTIMATE..........  3.482
```

VARIABLE NUMBER	REGRESSION COEFFICIENT	STD. ERROR OF REG. COEFF.	COMPUTED T-VALUE	DEGREES OF FREEDOM	LEVEL OF SIGNIFICANCE
1	3.94048	.53721	7.335	6	.00033
INTERCEPT	-112.86905				

FIGURE 27
COMPUTER OUTPUT
EXAMPLE 2

DEPENDENT -WEIGHT INDEPENDENT-HEIGHT AND AGE ---STEP-WISE MULTIPLE REGRESSION---

STEP 2

DEPENDENT VARIABLE............ 2

VARIABLE ENTERED..... 3

SUM OF SQUARES REDUCED IN THIS STEP.... 6.605
PROPORTION REDUCED IN THIS STEP........ .009

CUMULATIVE SUM OF SQUARES REDUCED...... 658.754
CUMULATIVE PROPORTION REDUCED.......... .909 OF 724.875

FOR 2 VARIABLES ENTERED

MULTIPLE CORRELATION COEFFICIENT... .953

F-VALUE FOR ANALYSIS OF VARIANCE... 24.907
DEGREES OF FREEDOM FOR AOV......... 2 AND 5

LEVEL OF SIGNIFICANCE OF F-VALUE FOR AOV .00251
STANDARD ERROR OF ESTIMATE......... 3.637

VARIABLE NUMBER	REGRESSION COEFFICIENT	STD. ERROR OF REG. COEFF.	COMPUTED T-VALUE	DEGREES OF FREEDOM	LEVEL OF SIGNIFICANCE
1	3.97540	.56330	7.057	5	.00088
3	.36670	.51887	.707	5	.51130
INTERCEPT	-124.25854				

FIGURE 28
COMPUTER OUTPUT
EXAMPLE 2

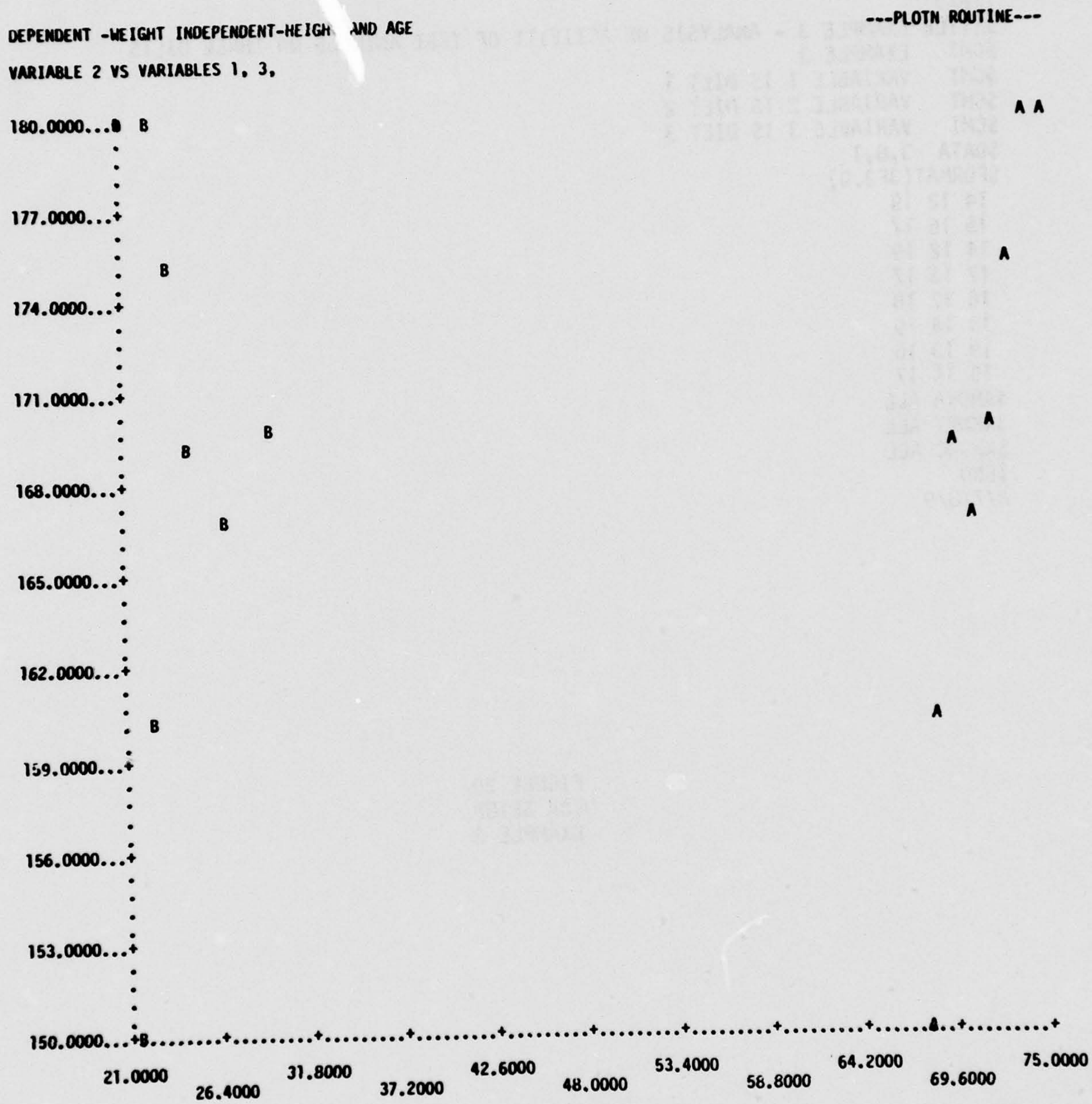


FIGURE 29
COMPUTER OUTPUT
EXAMPLE 2

```
JOB CARD
FETCHPS(GRASZ,GRAZZ,GRASS)
GRAZZ.
7/8/9
$TITLE EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS
$CMT   EXAMPLE 3
$CMT   VARIABLE 1 IS DIET 1
$CMT   VARIABLE 2 IS DIET 2
$CMT   VARIABLE 3 IS DIET 3
$DATA  3,8,1
$FORMAT(3F3.0)
 14 12 19
 15 16 17
 14 12 19
 17 13 17
 16 17 18
 18 14 19
 19 13 16
 15 16 17
$ANOVA ALL
$NKMRT ALL
$KRUWC ALL
$END
6/7/8/9
```

FIGURE 30
DECK SETUP
EXAMPLE 3

```
=========================================
*                                       *
*         GENERALIZED RESEARCH          *
*     ANALYSIS STATISTICAL SYSTEM       *
*                                       *
*   VERSION 4.00, LAST MOD 06 DEC 74    *
*                                       *
=========================================
```

UNLESS OTHERWISE NOTED, THE LEVEL OF SIGNIFICANCE PRINTED FOR

1. F OR CHI-SQUARE STATISTICS IS EQUAL TO THE PROBABILITY OF FINDING A RANDOM VALUE OF F OR CHI-SQUARE GREATER THAN THE OBTAINED VALUE OF F OR CHI-SQUARE, OR
2. T OR Z STATISTICS IS EQUAL TO TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T OR Z GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T OR Z VALUE.

FIGURE 31
COMPUTER OUTPUT
EXAMPLE 3

```
EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS

     *****  EXAMPLE 3

     *****  VARIABLE 1 IS DIET 1

     *****  VARIABLE 2 IS DIET 2

     *****  VARIABLE 3 IS DIET 3
```

FIGURE 32
COMPUTER OUTPUT
EXAMPLE 3

```
EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS

  RAW DATA DISPLAY---

1  1.40000000E+01  1.20000000E+01  1.90000000E+01

2  1.50000000E+01  1.60000000E+01  1.70000000E+01

3  1.40000000E+01  1.20000000E+01  1.90000000E+01

4  1.70000000E+01  1.30000000E+01  1.70000000E+01

5  1.60000000E+01  1.70000000E+01  1.80000000E+01

6  1.80000000E+01  1.40000000E+01  1.90000000E+01

7  1.90000000E+01  1.30000000E+01  1.60000000E+01

8  1.50000000E+01  1.60000000E+01  1.70000000E+01
```

FIGURE 33
COMPUTER OUTPUT
EXAMPLE 3

EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS

ONE WAY ANALYSIS OF VARIANCE 06 DEC 74

VARIABLE	NO. OF OBSERVATIONS	MEAN	STANDARD DEVIATION
1	8	16.0000	1.85164
2	8	14.1250	1.95941
3	8	17.7500	1.16496

OVERALL MEAN = 15.9583

ANALYSIS OF VARIANCE

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	P
TREATMENTS	52.58333	2	26.29167	9.1449	.00139
WITHIN TREATMENTS	60.37500	21	2.875000		
TOTAL	112.9583	23			

FIGURE 34
COMPUTER OUTPUT
EXAMPLE 3

EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS

---KEWMAN KUELS MULTIPLE RANGE TEST 06 DEC 74

VARIABLE	MEAN	NUMBER
1	1.60000000E+01	8
2	1.41250000E+01	8
3	1.77500000E+01	8

ERROR MEAN SQUARE = 2.87500000E+00

FIGURE 35
COMPUTER OUTPUT
EXAMPLE 3

```
EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS

---NEWMAN KUELS MULTIPLE RANGE TEST              06 DEC 74

MEANS IN ASCENDING ORDER

   2      1      3
      ----------------

---NOTE---  ANY TWO MEANS NOT UNDERSCORED BY THE SAME LINE
            ARE SIGNIFICANTLY DIFFERENT.

            ANY TWO MEANS UNDERSCORED BY THE SAME LINE
            ARE NOT SIGNIFICANTLY DIFFERENT.

       LEVEL OF SIGNIFICANCE = .05
```

FIGURE 36
COMPUTER OUTPUT
EXAMPLE 3

EXAMPLE 3 - ANALYSIS OF ACTIVITY OF TEST ANIMALS ON THREE DIETS

---THE KRUSKAL-WALLIS STATISTIC--- 06 DEC 74

	VAR. 1	RANK 1	VAR. 2	RANK 2	VAR. 3	RANK 3
1	1.4000E+01	6.000	1.2000E+01	1.500	1.9000E+01	22.500
2	1.5000E+01	8.500	1.6000E+01	11.500	1.7000E+01	16.000
3	1.4000E+01	6.000	1.2000E+01	1.500	1.9000E+01	22.500
4	1.7000E+01	16.000	1.3000E+01	3.500	1.7000E+01	16.000
5	1.6000E+01	11.500	1.7000E+01	16.000	1.8000E+01	19.500
6	1.8000E+01	19.500	1.4000E+01	6.000	1.9000E+01	22.500
7	1.9000E+01	22.500	1.3000E+01	3.500	1.6000E+01	11.500
8	1.5000E+01	8.500	1.6000E+01	11.500	1.7000E+01	16.000
RANK SUMS		98.500		55.000		146.500

THE STATISTIC,H= 10.49200
LEVEL OF SIGNIFICANCE = .0053

(MULTIPLE COMPARISONS)

(I, J)	AVERAGE RANK DIFFERENCE	Z(I,J)	LEVEL OF SIGNIFICANCE
(1, 2)	5.4375	1.5380	.062029
(1, 2)	6.0000	1.6971	.044843
(2, 3)	11.4375	3.2350	.000608

TO DECIDE WHICH PAIRS ARE DIFFERENT,SELECT A SIGNIFICANCE LEVEL A, FOR THE ENTIRE EXPERIMENT. THEN CALCULATE C=A/K(K-1). (K=NO. OF TREATMENTS) IF THE GIVEN SIGNIFICANCE LEVEL FOR COMPARISON I,J EXCEEDS C, DECIDE THAT THE PAIR DOES NOT DIFFER AT THE EXPERIMENTWISE ERROR LEVEL A.

FIGURE 37
COMPUTER OUTPUT
EXAMPLE 3

```
JOB CARD
FETCHPS (GRASZ,GRAZZ,GRASS)
GRAZZ.
7/8/9
$TITLE EXAMPLE  4, TWO STUDIES INVOLVING ACTIVITY OF SUBJECTS ON 2 DIETS
$CMT   EXAMPLE  4
$CMT   VARIABLE 1 IS CONTROL DIET STUDY 1
$CMT   VARIABLE 2 IS ENRICHED DIET STUDY 1
$CMT   VARIABLE 3 IS CONTROL DIET STUDY 2
$CMT   VARIABLE 4 IS ENRICHED DIET STUDY 2
$DATA  4,9,1
$FORMAT(4F3.0)
 15 13 13 15
 17 19 15 19
 10 15 14 13
 19 17 10 20
 12 16 12 17
 13 15 17 16
 16 18 17 15
       18 18
       15
$BSTAT ALL
$TITLE (EXAMPLE 4) PAIRED T-TEST FOR STUDY 1 DATA
$PTST  1,2
$TITLE (EXAMPLE 4) WILCOXON'S SIGNED RANK TEST FOR STUDY 1 DATA
$SNRNK 1,2
$TITLE (EXAMPLE 4) NON-PAIRED T-TEST FOR STUDY 2 DATA
$NTST  3,4
$TITLE (EXAMPLE 4) WILCOXON'S RANK SUM TEST FOR STUDY 2 DATA
$RNKSM 3,4
$END
6/7/8/9
```

FIGURE 38
DECK SETUP
EXAMPLE 4

```
****************************************
*                                      *
*        GENERALIZED RESEARCH          *
*    ANALYSIS STATISTICAL SYSTEM       *
*                                      *
*  VERSION 4.00, LAST MOD 06 DEC 74    *
*                                      *
****************************************
```

UNLESS OTHERWISE NOTED, THE LEVEL OF SIGNIFICANCE PRINTED FOR

1. F OR CHI-SQUARE STATISTICS IS EQUAL TO THE PROBABILITY OF FINDING A RANDOM VALUE OF F OR CHI- SQUARE GREATER THAN THE OBTAINED VALUE OF F OR CHI-SQUARE, OR
2. T OR Z STATISTICS IS EQUAL TO TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T OR Z GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T OR Z VALUE.

FIGURE 39
COMPUTER OUTPUT
EXAMPLE 4

```
EXAMPLE 4,TWO STUDIES INVOLVING ACTIVITY OF SUBJECTS ON 2 DIETS

   *****  EXAMPLE 4

   *****  VARIABLE 1 IS CONTROL DIET STUDY 1

   *****  VARIABLE 2 IS ENRICHED DIET STUDY 1

   *****  VARIABLE 3 IS CONTROL DIET STUDY 2

   *****  VARIABLE 4 IS ENRICHED DIET STUDY 2
```

FIGURE 40
COMPUTER OUTPUT
EXAMPLE 4

EXAMPLE 4,TWO STUDIES INVOLVING ACTIVITY OF SUBJECTS ON 2 DIETS

RAW DATA DISPLAY---

1	1.50000000E+01	1.30000000E+01	1.30000000E+01	1.50000000E+01
2	1.70000000E+01	1.90000000E+01	1.50000000E+01	1.90000000E+01
3	1.00000000E+01	1.50000000E+01	1.40000000E+01	1.30000000E+01
4	1.90000000E+01	1.70000000E+01	1.00000000E+01	2.00000000E+01
5	1.20000000E+01	1.60000000E+01	1.20000000E+01	1.70000000E+01
6	1.30000000E+01	1.50000000E+01	1.70000000E+01	1.60000000E+01
7	1.60000000E+01	1.80000000E+01	1.70000000E+01	1.50000000E+01
8	-0.	-0.	1.80000000E+01	1.80000000E+01
9	-0.	-0.	1.50000000E+01	-0.

FIGURE 41
COMPUTER OUTPUT
EXAMPLE 4

EXAMPLE 4,TWO STUDIES INVOLVING ACTIVITY OF SUBJECTS ON 2 DIETS

---BSTAT ROUTINE---

VARIABLE	OBSERVATIONS	MINIMUM	MAXIMUM	RANGE	MEDIAN
1	7	1.00000000E+01	1.90000000E+01	7.00000000E+00	1.50000000E+01
2	7	1.30000000E+01	1.90000000E+01	7.00000000E+00	1.60000000E+01
3	9	1.00000000E+01	1.80000000E+01	7.00000000E+00	1.50000000E+01
4	8	1.30000000E+01	2.00000000E+01	7.00000000E+00	1.65000000E+01

VARIABLE	MEAN	ST. DEVIATION	ST.ERR.OF MEAN
1	1.45714286E+01	3.10145895E+00	1.17224130E+00
2	1.61428571E+01	2.03540098E+00	7.69309258E-01
3	1.45555556E+01	2.60341656E+00	8.67805520E-01
4	1.66250000E+01	2.32609421E+00	8.22398496E-01

FIGURE 42
COMPUTER OUTPUT
EXAMPLE 4

(EXAMPLE 4)PAIRED T-TEST FOR STUDY 1 DATA

---PAIRED T-TEST ROUTINE--- 06 DEC 74

VARIABLE	MEAN	STANDARD DEVIATION OF DIFFERENCES	T-VALUE	DEGREES OF FREEDOM
1	1.45714286E+01			
VERSUS				
2	1.61428571E+01	2.69920623E+00	-1.54030809E+00	6

LEVEL OF SIGNIFICANCE = .17441517

FIGURE 43
COMPUTER OUTPUT
EXAMPLE 4

```
(EXAMPLE4)WILCOXON‡S SIGNED RANK TEST FOR STUDY 1 DATA

---WILCOXON‡S SIGNED RANK TEST---

VARIABLE  1 IS X.  VARIABLE  2 IS Y.

           X          Y   DIFFERENCE        SIGN        RANK
     15.0000    13.0000       2.0000           +      3.0000
     17.0000    19.0000      -2.0000           -      3.0000
     10.0000    15.0000      -5.0000           -      7.0000
     19.0000    17.0000       2.0000           +      3.0000
     12.0000    16.0000      -4.0000           -      6.0000
     13.0000    15.0000      -2.0000           -      3.0000
     16.0000    18.0000      -2.0000           -      3.0000

   NO. OF NON-ZERO DIFFERENCES =             7


THE SUM OF THE + RANKS IS      6.0000
THE SUM OF THE - RANKS IS     22.0000
  THE MINIMUM OF THESE IS      6.0000

LEVEL OF SIGNIFICANCE =         .10938
(THIS IS THE PROBABILITY OF FINDING A RANDOM
VALUE ≤ THE MINIMUM SUM.  IF THE MINIMUM IS NOT A
WHOLE NUMBER, THE SIGNIFICANCE LEVEL LEVEL IS THE
PROBABILITY OF FINDING A RANDOM VALUE < THE NEXT
HIGHEST WHOLE NUMBER.)

LARGE SAMPLE APPROXIMATION,AT  =     -1.4033

UPPER TAIL PROBABILITY FOR ABS(AT) =     .0803
```

FIGURE 44
COMPUTER OUTPUT
EXAMPLE 4

```
(EXAMPLE 4)NON-PAIRED T-TEST FOR STUDY 2 DATA

--- NON-PAIRED T-TEST ---          06 DEC 74

VARIABLE        MEAN          STAND. DEVIATION   OBSERVATIONS
   3      1.45555556E+01      2.60341656E+00          9
VERSUS
   4      1.66250000E+01      2.32609421E+00          8

T=  -1.71877063E+00    DEGREES OF FREEDOM =    15

LEVEL OF SIGNIFICANCE =              .10621956
```

FIGURE 45
COMPUTER OUTPUT
EXAMPLE 4

(EXAMPLE 4)WILCOXON S RANK SUM TEST FOR STUDY 2 DATA 06 DEC 74

---WILCOXON S RANK SUM TEST---

VARIABLE 3 IS X. VARIABLE 4 IS Y.

	X	Y	RANK OF Y
1	13.0000	15.0000	7.5000
2	15.0000	19.0000	16.0000
3	14.0000	13.0000	3.5000
4	10.0000	20.0000	17.0000
5	12.0000	17.0000	12.0000
6	17.0000	16.0000	10.0000
7	17.0000	15.0000	7.5000
8	18.0000	18.0000	14.5000
9	15.0000		
		W =	88.0000

UPPER TAIL PROBABILITY FOR W= .0697362794
UPPER TAIL PROBABILITY FOR N(M+N+1)-W= .9476610581

LARGE SAMPLE APPROXIMATION,AW,IS 1.5549

THE PROBABILITY,P,OF A RANDOM
VALUE FROM THE STANDARD NORMAL
DISTRIBUTION BEING GREATER THAN AW IS .0600

*NOTE-FOR NEGATIVE AW USE 1-P

FIGURE 46
COMPUTER OUTPUT
EXAMPLE 4

---SIEGEL-TUKEY TEST---

	X	Y	RANK OF Y
1	13.0000	15.0000	14.5000
2	15.0000	19.0000	3.0000
3	14.0000	13.0000	6.5000
4	10.0000	20.0000	2.0000
5	12.0000	17.0000	11.6667
6	17.0000	16.0000	15.0000
7	17.0000	15.0000	14.5000
8	18.0000	18.0000	6.5000
9	15.0000		

		W =	73.6667

UPPER TAIL PROBABILITY FOR W= .4463617930
UPPER TAIL PROBABILITY FOR N(M+N+1)-W= .5955073153

FIGURE 47
COMPUTER OUTPUT
EXAMPLE 4

```
JOB CARD
FETCHPS (GRASZ,GRAZZ,GRASS)
GRAZZ.
7/8/9
$TITLE EXAMPLE 5A,USING $TRANS FUNCTIONS AND ARITHMETIC FUNCTIONS
$DATA  2,4,1
$TRANS X(3)=SIN(X(1))
$TRANS X(4)=ASIN(X(3))
$TRANS X(5)=LOG(X(2))
$TRANS   X(6)=10*(X(1)↑2+X(5))
$FORMAT(2F5.0)
    0   15
  .87   25
1.239   30
1.501   90
$TITLE EXAMPLE   5B,USING $TRANS IF-THEN-ELSE STATEMENTS
$CMT   THE FIRST 3$TRANS STATEMENTS SHOW HOW TO RECODE BLANK DATA
$DATA  3,9,1
$TRANS IF X(3)=0 THEN IF LOGICAL (X(3))=1 THEN X(4)=2
$TRANS ELSE X(4)=X(3)
$TRANS ELSE X(4)=X(3)
$TRANS IF X(1)=2 THEN X(5)=X(4)↑2 ELSE X(5)=X(2)+X(1)
$FORMAT(3F5.0)
    0    0    1
    1    1
    2    1
    3    2    4
    4    2    5
    5    3    6
         3    7
         4    8
         4    9
$TITLE EXAMPLE 5C, USING $TRANS TRANSPOSE AND $ADD
$DATA  6,2,1
$FORMAT(6F2.0)
 1 2 3
 4 5 6 7 8 9
$TRANS TRANSPOSE
$EXECU
$ADD   1,2,3
$END
6/7/8/9
```

FIGURE 48
DECK SETUP
EXAMPLE 5

```
==========================================
*                                        *
*          GENERALIZED RESEARCH          *
*      ANALYSIS STATISTICAL SYSTEM       *
*                                        *
*   VERSION 4.00, LAST MOD 05 DEC 74     *
*                                        *
==========================================
```

UNLESS OTHERWISE NOTED, THE LEVEL OF SIGNIFICANCE PRINTED FOR

1. F OR CHI-SQUARE STATISTICS IS EQUAL TO THE PROBABILITY OF FINDING A RANDOM VALUE OF F OR CHI-SQUARE GREATER THAN THE OBTAINED VALUE OF F OR CHI-SQUARE, OR
2. T OR Z STATISTICS IS EQUAL TO TWICE THE PROBABILITY OF FINDING A RANDOM VALUE OF T OR Z GREATER THAN THE ABSOLUTE VALUE OF THE OBTAINED T OR Z VALUE.

FIGURE 49
COMPUTER OUTPUT
EXAMPLE 5

```
EXAMPLE 5A,USING $TRANS FUNCTIONS AND ARITHMETIC FUNCTIONS

    *****  X(3)=SIN(X(1))

    *****  X(4)=ASIN(X(3))

    *****  X(5)=LOG(X(2))

    *****   X(6)=10*(X(1)↑2+X(5))
```

FIGURE 50
COMPUTER OUTPUT
EXAMPLE 5

EXAMPLE 5A,USING $TRANS FUNCTIONS AND ARITHMETIC FUNCTIONS

RAW DATA DISPLAY---

1	0.	1.50000000E+01	0.	0.	1.17609126E+00	1.17609126E+01
2	8.70000000E-01	2.50000000E+01	7.64328937E-01	8.70000000E-01	1.39794001E+00	2.15484001E+01
3	1.23900000E+00	3.00000000E+01	9.45458730E-01	1.23900000E+00	1.47712125E+00	3.01224225E+01
4	1.50100000E+00	9.00000000E+01	9.97565225E-01	1.50100000E+00	1.95424251E+00	4.20724351E+01

FIGURE 51
COMPUTER OUTPUT
EXAMPLE 5

```
EXAMPLE 5B,USING $TRANS IF-THEN-ELSE STATEMENTS

*****     THE FIRST 3 $TRANS STATEMENTS SHOW HOW TO RECODE BLANK DATA
```

FIGURE 52
COMPUTER OUTPUT
EXAMPLE 5

EXAMPLE 5B,USING $TRANS IF-THEN-ELSE STATEMENTS

```
*****  IF X(3)=0 THEN IF LOGICAL (X(3))=1 THEN X(4)=2
*****  ELSE X(4)=X(3)
*****  ELSE X(4)=X(3)
*****  IF X(1)=2 THEN X(5)=X(4)↑2  ELSE X(5)=X(1)
```

FIGURE 53
COMPUTER OUTPUT
EXAMPLE 5

```
EXAMPLE 5B,USING $TRANS IF-THEN-ELSE STATEMENTS

---RAW DATA DISPLAY---

1  0.              0.              1.00000000E+00  1.00000000E+00  0.
2  1.00000000E+00  1.00000000E+00  -0.             2.00000000E+00  2.00000000E+00
3  2.00000000E+00  1.00000000E+00  -0.             2.00000000E+00  4.00000000E+00
4  3.00000000E+00  2.00000000E+00  4.00000000E+00  4.00000000E+00  5.00000000E+00
5  4.00000000E+00  2.00000000E+00  5.00000000E+00  5.00000000E+00  6.00000000E+00
6  5.00000000E+00  3.00000000E+00  6.00000000E+00  6.00000000E+00  8.00000000E+00
7 -0.              3.00000000E+00  7.00000000E+00  7.00000000E+00  -0.
8 -0.              4.00000000E+00  8.00000000E+00  8.00000000E+00  -0.
9 -0.              4.00000000E+00  9.00000000E+00  9.00000000E+00  -0.
```

FIGURE 54
COMPUTER OUTPUT
EXAMPLE 5

```
EXAMPLE 5C, USING $TRANS TRANSPOSE AND $ADD

  ---RAW DATA DISPLAY---

1  1.00000000E+00   2.00000000E+00   3.00000000E+00  -0.              -0.              -0.
2  4.00000000E+00   5.00000000E+00   6.00000000E+00   7.00000000E+00   8.00000000E+00   9.00000000E+00
```

FIGURE 55
COMPUTER OUTPUT
EXAMPLE 5

```
EXAMPLE 5C, USING $TRANS TRANSPOSE AND $ADD

      *****  TRANSPOSE

1  1.00000000E+00   4.00000000E+00

2  2.00000000E+00   5.00000000E+00

3  3.00000000E+00   6.00000000E+00

4 -0.               7.00000000E+00

5 -0.               8.00000000E+00

6 -0.               9.00000000E+00

      *****     EXECU
```

FIGURE 56
COMPUTER OUTPUT
EXAMPLE 5

EXAMPLE 5C, USING $TRANS TRANSPOSE AND $ADD

---RAW DATA DISPLAY---

VARIABLES 1 AND 2HAVE BEEN JOINED IN VARIABLE 3

1	1.00000000E+00	4.00000000E+00	1.00000000E+00
2	2.00000000E+00	5.00000000E+00	2.00000000E+00
3	3.00000000E+00	6.00000000E+00	3.00000000E+00
4	-0.	7.00000000E+00	4.00000000E+00
5	-0.	8.00000000E+00	5.00000000E+00
6	-0.	9.00000000E+00	6.00000000E+00
10	-0.	-0.	7.00000000E+00
11	-0.	-0.	8.00000000E+00
12	-0.	-0.	9.00000000E+00

FIGURE 57
COMPUTER OUTPUT
EXAMPLE 5

DISTRIBUTION LIST

US Army Medical Research and Development Command Washington, DC 20314	5
Defense Documentation Center ATTN: DDC-TCA Alexandria, VA 22314	20
Superintendent Academy of Health Sciences, US Army ATTN: AHS-COM Ft Sam Houston, TX 78234	1
Dir of Biol & Med Sciences Div Office of Naval Research 800 N. Quincy Street Arlington, VA 22217	1
CO, Naval Medical R&D Command National Naval Medical Center Bethesda, MD 20014	1
Dir of Prof Svcs Office of the Surgeon General Department of the Air Force Washington, DC 20314	1
Dir of Defense Research and Engineering ATTN: Asst Dir (Environmental and Life Sciences) Washington, DC 20301	1