JUL 19 1967

MASTER

OAK RIDGE NATIONAL LABORATORY
operated by
UNION CARBIDE CORPORATION
for the
U.S. ATOMIC ENERGY COMMISSION

ORNL - TM - 1863

10

STATPAK

Norma C. Hull

NOTICE This document contains information of a preliminary nature and was prepared primarily for internal use at the Oak Ridge National Laboratory. It is subject to revision or correction and therefore does not represent a final report.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

- A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
- B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

ORNL-TM-1863

Contract No. W-7405-eng-26

MATHEMATICS DIVISION

CESTI PRICES

STATPAK

H.C. \$ 3.00, MN .65

Norma C. Hull

JULY 1967

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee
operated by
UNION CARBIDE CORPORATION
for the
U.A. ATOMIC ENERGY COMMISSION

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK

STATPAK

Norma C. Hull, Editor

Abstract

This memo describes three classes of statistical computer programs which are available for use on the ORNL computers.

The classes are:

- (1) regression analysis, linear and non-linear;
- (2) analysis of variance, for one-way classifications, for factorial experiments, and for three way classifications with multiple observations;
- (3) graphing programs.

Some programs are modifications of the BMD and BMD series written at the University of California at Los Angeles; some are original programs developed by the Statistics Section of the Mathematics Division; and others are modifications of programs prepared by personnel of other divisions of ORNL.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

**THIS PAGE
WAS INTENTIONALLY
LEFT BLANK**

Acknowledgments

The guidance and assistance of the Statistics Section of the Mathematics Division, and in particular Drs. M. A. Kastenbaum, D. A. Gardiner, and D. G. Gosslee and Mr. T. L. Hebble in the development and modification of the programs are gratefully acknowledged.

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK

TABLE OF CONTENTS

	Page
Regression Analysis	
EIMD29GR	1
BIMD31GR	20
GENLIN	41
NONLIN	68
Analysis of Variance	
ANOVA	82
ANOVAGR	88
BMD02VH	103
TWAOV	128
YATES	136
Graphical Analysis	
GRAPH	147
ELLIPSE	153
Appendix A	162
Appendix B	164
Appendix C	166

BIMD29GR
Norma C. Hull
(Fortran 63)

DESCRIPTION

This program is a modification of the program "BIMD-29" written at the University of California at Los Angeles. BIMD29GR is a multiple regression and correlation program with the maximum number of variables which can be processed limited to 60 variables. From the several variables, one variable is selected as the dependent variable and one or more variables are selected as independent variables. This is accomplished by designating the dependent variable by number and by deleting the variables which are to be ignored. The Replacement and Deletion Card(s), which are discussed in more detail on following pages, sets up the model for each problem. There is no limit to the number of models which can be constructed. The maximum number of variables which can be deleted at one time, however, is 31. The upper limitation on sample size without using the graphing option for a specific fit is 99,999, but the upper limit on sample size is 275 when graphs are specified. The sample size must be at least as large as the number of variables plus two. This program can perform transformations on any variable and generate new variables if desired, according to the codes specified on the transgeneration cards. Any number of variables can be generated, but the total number of variables must not exceed 60. Any variable, original or generated, can be named the dependent or independent variables. It is assumed that the dependent variable, Y,

is normally distributed with mean μ and variance σ^2 , where

$$\mu = A + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p.$$

Input includes the dependent variable and any independent variables observed, the number of original variables including the dependent variable, sample size, and the number of models to be constructed by the Replacement and Deletion cards.

The output includes a list of the regression coefficients in the order in which they would be chosen by a "stepwise regression" algorithm.

Additional output includes:

- (1) Standard errors of estimate,
- (2) Sums and sums of squares,
- (3) Cross product sums
- (4) Cross products of deviations,
- (5) Simple correlation coefficients,
- (6) Proportions of variance
- (7) F-values,
- (8) F-values,
- (9) Graphs of residuals against any of the variables,
- (10) Critical values of Student's t at the 10, 5 and 1% levels.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ EXECUTE Card
- (3) Binary Program Deck
- (4) Δ Card
- (5) Graphing Option Card
- (6) Problem Card for First Problem
- (7) Transgeneration Card(s) (if any)
- (8) Variable Name Card(s)

- (9) Scale Card(s)
- (10) Variable Format Card(s)
- (11) Input Data Cards
- (12) R-D Card(s)
- (13) Graph Card (if used)

Cards (12) and (13) may be repeated for each problem, and cards 6-13 may be repeated for as many problems as desired.

Master Control System Card

Starting in column 1 the following characters appear:

Δ COOP, Request Number, Name, Ø/17, G.

Graphing Option Card

If graphs are desired for any fit in any problem, a "1" must be punched in column one. If no graphs are desired, a "0" (zero) must be punched in column one. This card sets up the plotting tape. It must precede the Problem Card for the first problem but for none of the subsequent problems.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	PR	Problem
3-6		Problem number
7,8		Number of original variables, p, ($p \leq 60$)
9-13		Sample size, n
14-15	00	No transgeneration
		m transgeneration cards ($M \leq 99$)

16,17	00	No variable added to original set after transgeneration q variables added to total set after transgeneration
18-20		Number of Replacement and Deletion cards
21-25		Per cent of sum of squares to limit variables entering in a regression model. Key punch a value with a decimal point. (See Note 1 below)
26-76		Alphanumeric characters to form a title for identification of problem
71,72	00	No variable format cards (input data <u>must</u> be punched according to a (12F6.0) format when not using a variable format card k variable format cards ($K \leq 5$)

Note: (1) The choice of percentage value for limiting variable will depend upon the purpose of the analysis. A suggested trial value is one percent, 0.01. If all independent variables are to be included in a regression model, keypunch 0.0.

Transgeneration Card(s)

If a non-zero number is specified in column 14 and 15 of the Problem Card, the same number of Transgeneration Card(s) must be prepared. Different types of transformation can be performed successively on the same variable, if desired. See Appendix A for an explanation of transgeneration. The format of the Transgeneration Card is as follows:

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	TR	TRansgeneration Card
3,4		The number to be assigned to the transformed or generated variable

5,6

Transformation code
(Codes 01, 02, ..., 14, 16, ..., 27 of the
transgeneration list, Appendix A)*

7,8

The number of the variable to be transgenerated
or A-variable

9-14

Constant or B-variable (keypunch a constant with
a decimal point; constant may be negative.)

*Code 15 of transgeneration list Appendix A may not be used with BMD29GR.

Variable Name Card(s)

The purpose of Variable Name Card(s) is to identify the variables
in the output by their names. The format of Variable Name Card is as
follows:

<u>Column</u>	<u>Explanation</u>
1-6	Name of 1st Variable
7-12	" " 2nd "
13-18	" " 3rd "
⋮	
67-72	" " 12th "

If there are more than 12 variables, continue keypunching on the
second card in the same manner. If variables are to be generated,
the names of these variables must also be keypunched in addition to
those of the original variables. If the identification of the vari-
ables by their names is not desired, use blank card(s) for this purpose.
Variable Name Card(s) must be present.

Scale Card(s)

The numbers punched in the Scale Card(s) are signed integers (+ need not be punched) which indicate the power of ten by which the input variable is to be multiplied. If no scaling is desired, blank fields must be indicated for each input variable. The fields are right justified, and the format of this card is I6.

<u>Columns</u>	<u>Explanation</u>
1-6	Scale factor for first variable
7-12	" " " second "
⋮	
67-72	" " " twelfth "

Scaling continues on the next card if there are more than twelve variables. Scale Card(s) must be present.

Format Card(s)

Columns 1-72 of as many as five cards may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats.

Input Data Cards

The format for reading data in this program is variable. The data must be input in the same format as specified on the Format Card(s).

Replacement and Deletion Card(s)

This card, R-D card, has a fivefold purpose:

- (1) It indicates the replacement of the dependent variable
- (2) It indicates the deletion of variables
- (3) It indicates a replacement and deletion
- (4) It controls the output from the model fitted, and
- (5) It sets up the graphing subroutine.

The replacement of a dependent variable and deletion of different sets of variables can be made as many times as desired. This means that it is possible to make multiple regression and correlation analyses with selected sets of variables. The program will retain all the variables, original and generated, until all R-D Cards for one problem are processed. R-D Card(s) must be present, and the number of R-D Card(s) prepared must agree with the number specified in col. 18-20 of the Problem Card.

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	RD	(for Replacement and Deletion Card)
3,4	02	If the tables of residuals and analysis of extreme residuals are desired.
	01	If only the analysis of extreme residuals is desired
	00	If neither one is desired
5,6		A variable number to be treated as the next dependent variable (ND)
7,8		Total number of variables to be deleted (N2)
9,10		Number of graphs for the specific R-D fit (IXP)
11,12		1st variable to be deleted
13,14		2nd " " " "
⋮		
71,72		31st variable to be deleted.

When the sample size is large, 02 punch in columns 3,4 is not recommended.

The maximum number of variables which can be deleted at one time is 31.

Although Replacements and Deletions are repeated over and over again, the arrangement of variables will not change. Therefore, the number of the next dependent variable and the numbers of variables to be deleted must be stated in terms of the basic set of variables, which may be the original set of variables after any transgenerations are performed.

Delete low numbered variables first; namely X_2 comes first and then X_7 .

Graph Card (if used)

If IXP is a positive number in columns 9,10 on the R-D Card, variable numbers must be specified to designate the independent variable to be used in the procedure. The dependent variable is always the residual of the fit from the preceding R-D card. The total number of independent variables must agree with IXP. The format for the Graph Card is I2 starting in column one. The number of graphs per R-D card is limited to 10.

EXAMPLE

In the example for BIND29GR a trigonometric regression model is desired. The model is

$$Y = A+Bx+C \sin(\pi/4X)+D \cos(\pi/4X)+E \sin(\pi X)+F \cos(\pi X)$$

modulo 2π . Modulo 2π is obtained by using transgenerations. The first RD card indicates that two variables (one and four or "X" and "Sin (πX)") are not to be included in the fit. The second RD Card omits variable four or "Sin(πX)" from the fit. A graph of the residuals is desired for the second fit.

The 1604A job card must show that tape drive "17" is used for plotting.

|
 PH 2 1015 4 2 0.0 $Y = A + BX + CSINP/4 * X + DCOS P/4 * X + ESINP * X + FCOS P * X$
 TR0608 2 0.0
 TR0209 13.1416
 TR0209 2.25
 TR0224 26.2832
 TR0216 2
 TR0309 13.1416
 TR0309 3.25
 TR0324 36.2832
 TR0317 3
 TR0409 13.1416
 TR0424 46.2832
 TR0416 4
 TR0509 13.1416
 TR0524 56.2832
 TR0517 5
 X S P/4XC P/4X S PX C PX Y

(2F10.0)

1	797
2	805
3	804
4	806
5	780
6	787
7	776
8	791
9	783
10	798

RD02 0 2 0 1 4

RD02 0 1 1 4

|

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

MULTIPLE CORRELATION AND REGRESSION - PROBLEM $Y = A + BX + CSINP/4 \cdot X + DCOSP/4 \cdot X + ESINP \cdot X + FCOSP \cdot X$

NO. OF VARIABLES 6

SAMPLE SIZE 10

TRANSFORMATION CARD NO.	VARIABLE NO. ASSIGNED	TYPE OF TRANSFORMATION	VARIABLE TRANSFORMED	2ND VARIABLE OR CONSTANT
1	6	8	2	0
2	2	9	1	3.14159999997+00
3	2	9	2	2.49999999994+01
4	2	24	2	6
5	2	16	2	0
6	3	9	1	3.14159999997+00
7	3	9	3	2.49999999994+01
8	3	24	3	6
9	3	17	3	0
10	4	9	1	3.14159999997+00
11	4	24	4	6
12	4	16	4	0
13	5	9	1	3.14159999997+00
14	5	24	5	6
15	5	17	5	0

SUMS

5.9000000000+01 1.7071007335+00 7.0711954513+01 -3.6732151616+05 0 7.9269999999+03

SUMS OF SQUARES

3.8500000000+02 5.4999944898+00 4.5000055098+00 2.6985019262+10 1.0000000000+01 6.2848049999+06

CROSS PRODUCT SUMS

ROW 1

3.8500000000+02 6.7071098629+00 1.0364054496+01 -1.8366029242+04 5.0000000000+00 4.3452000000+04

ROW 2

6.7071098629+00 5.4999944898+00 4.9998850025+01 -5.1941934041+06 2.9288457383+01 1.4014794027+03

ROW 3

1.0364054496+01 4.9998850025+01 4.5000055098+00 -5.1945799236+06 -7.0711219874+01 5.3089615429+02

ROW 4

-1.8366029242+04 -5.1941934041+06 -5.1945799236+06 2.6985019262+10 3.6732151617+05 -2.8944931981+02

ROW 5

5.0000000000+00 2.9288457383+01 -7.0711219874+01 3.6732151617+05 1.0000000000+01 4.7000000000+01

ROW 6

4.3452000000+04 1.4014794027+03 5.3089615429+02 -2.8944931981+02 4.7000000000+01 6.2848049999+06

CROSS PRODUCTS OF DEVIATIONS

ROW 1

8.2500000000+01 -2.6819441710+00 4.4748969984+00 1.8366541468+05 5.0000000000+00 -1.4650000000+02

ROW 2

-2.6819441710+00 9.2085751984+00 5.7927655083+01 1.0763548925+06 2.9288457383+01 -4.8260651260+01

ROW 3

PROBLEM NUMBER

REPLACEMENT AND DELETION 1

DEPENDENT VARIABLE IS NOW 6

NUMBER OF VARIABLES DELETED 2

VARIABLES DELETED... 1 4

ANALYSIS OF VARIANCE FOR REGRESSION

SOURCE OF VARIATION	D.F.	SUM OF SQUARES	MEAN SQUARES	F VALUE
DUE TO REGRESSION	3	8.485097782**02	2.828365927**02	7.589861498**00
DEVIATION ABOUT REGRESSION	6	2.235903194**02	3.726505323**01	
TOTAL	9	1.072100098**03		

COVARIANCE MATRIX (INVERSE MATRIX TIMES VARIANCE) IN ORDER OF ENTRY INTO REGRESSION EQUATION

INTERCEPT	3.9636251212**00	-1.1855938461**00	-4.9110487392**01	-2.4098882638**06
VARIABLE 2	-1.1855938461**00	7.2163244028**00	-6.5605281822**01	-2.5775994904**01
VARIABLE 3	-4.9110487392**01	-6.5605281822**01	8.5289637772**00	6.2230820796**01
VARIABLE 5	-2.4098882638**06	-2.5775994904**01	6.2230820796**01	3.7780538869**00

INTERCEPT (A-VALUE) IS 7.9155516810**02 WITH STANDARD ERROR 1.998855118**00

NO.	NAME	MEAN	STANDARD DEVIATION	REGRESSION COEFFICIENT	STD. ERROR OF REG. COEF.	COMPUTED T VALUES	PARTIAL CORR. COEF.
2	S P/4Y	1.707100733**01	7.607434374**01	6.542227453**00	2.686414786**00	3.552201143**00	8.232524501**01
3	C P/4Y	7.07115451**02	7.031677297**01	-6.847978742**00	2.928438970**00	-2.344845694**00	-6.615083454**01
5	C PX	0	1.054092553**00	3.836273445**00	1.943722945**00	2.025120635**00	6.371860683**01

DEPENDENT

6 Y 7.92700000**02 1.09143132**01

COMPARE CHECK ON FINAL COEFFICIENT 3.936274445**06 T(1)= 1.94317**00 T(05)= 2.44691**00 T(01)= 3.70742**00

INCREMENTS FOR INDEPENDENT VARIABLES

VARIABLE	SUMS OF SQUARES	PROP. VAR.	F VALUE
NO. NAME			
2 S P/4Y	4.471646029**02	4.170922135**01	5.724297712**00
3 C P/4Y	2.485169592**02	2.318038771**01	4.621501202**00
5 C PX	1.528282162**02	1.425503239**01	4.101113588**00

CUMULATIVE REGRESSIONS

VARIABLE	STD. ERROR OF ESTIMATE	SUMS OF SQUARES	PROP. VAR.	F VALUE	MULTIPLE R
NO. NAME					

2	S P/4X	8.838378632**00	4.471646029**02	4.170922135**01	5.724297712**00	6.458257674**01
3	C P/4X	7.333080970**00	6.556815621**02	5.488960906**01	6.468548897**00	8.055458683**01
5	C PX	6.104510893**00	8.485097782**02	7.914464145**01	7.589361498**00	8.896327413**01

PROPORTION OF VARIANCE SPECIFIED TO LIMIT VARIABLES

0

TABLE OF RESIDUALS

OBSERVATION	Y VALUE	Y ESTIMATE	RESIDUAL
1	7.970000000**02	7.895245245**02	7.475475535**00
2	8.050000000**02	8.050343862**02	*3.438615799**02
3	8.040000000**02	7.992089971**02	4.791002914**00
4	8.060000000**02	8.023393422**02	3.660657823**00
5	7.800000000**02	7.857131637**02	*5.713163733**00
6	7.870000000**02	7.859484306**02	1.051569372**00
7	7.760000000**02	7.760287902**02	*2.879025042**02
8	7.910000000**02	7.886434548**02	2.356545195**00
9	7.830000000**02	7.895245244**02	*6.524524450**00
10	7.980000000**02	8.050343862**02	*7.034386158**00

RATIO OF RANGES FOR THE SMALLEST RESIDUAL 4.311585053**02

RATIO OF RANGES FOR THE LARGEST RESIDUAL 1.917480445**01

CRITICAL VALUE OF THE RATIO AT ALPHA = .10 4.090000000**01

PROBLEM NUMBER:
REPLACEMENT AND DELETION 2

DEPENDENT VARIABLE IS NOW 6
NUMBER OF VARIABLES DELETED 1
VARIABLES DELETED... 4

ANALYSIS OF VARIANCE FOR REGRESSION

SOURCE OF VARIATION	D.F.	SUM OF SQUARES	MEAN SQUARES	F VALUE
DUE TO REGRESSION	4	9.865693260**02	2.466423315**02	1.441833897**01
DEVIATION ABOUT REGRESSION	5	8.55377161**01	1.710615432**01	
TOTAL	9	1.072100098**03		

COVARIANCE MATRIX (INVERSE MATRIX TIMES VARIANCE) IN ORDER OF ENTRY INTO REGRESSION EQUATION

INTERCEPT	9.5313696503**00	*1.4809345761**00	2.0357553462**00	8.8840511136**01	*1.4021660142**00
VARIABLE 2	*1.4809345761**00	3.4265604937**00	*5.7577294178**01	*2.2621775198**01	1.7029094781**01
VARIABLE 3	2.0357553462**00	*5.7577294178**01	4.5781359858**00	5.4615217696**01	*4.1112619067**01
VARIABLE 5	8.8840511136**01	*2.2621775198**01	5.4615217696**01	1.8366243550**00	*1.6152853640**01
VARIABLE 1	*1.4021660142**00	1.7029094781**01	*4.1112619067**01	*1.6152853640**01	2.5493948555**01

INTERCEPT (A VALUE) IS 7.9944444998**02 WITH STANDARD ERROR 3.0372916367**00

VARIABLE	MEAN	STANDARD DEVIATION	REGRESSION COEFFICIENT	STD. ERROR OF REG. COEF.	COMPUTED T VALUES	PARTIAL CORR. COEF.
NO. NAME						
2 S P/4X	1.707100733**01	7.607434374**01	8.584785093**00	1.851097106**00	4.637674094**00	9.007651452**01
3 C P/4X	7.071195451**02	7.031677297**01	*4.534776407**00	2.139567914**00	*2.119393188**00	*6.879179093**01
5 C PX	0	1.054092553**00	4.845114079**00	1.355221147**00	3.575146455**00	8.478275752**01
DEPENDENT	5.500000000**00	3.027650354**00	*1.434417526**00	5.049153252**01	*2.840907088**00	*7.857904115**01
6 Y	7.927000000**02	1.09143132**01				

COMPARE CHECK ON FINAL COEFFICIENT = 1.434417526**00 T(-) = -2.01499**00 T(.05) = -2.57054**00 T(.01) = -4.03224**00

INCREMENTS FOR INDEPENDENT VARIABLES

VARIABLE	SUMS OF SQUARES	PROP. VAR.	F VALUE EACH TERM
NO. NAME			
2 S P/4X	4.471646029**02	4.170922135**01	5.724297712**00
3 C P/4X	2.485169592**02	2.318038771**01	4.621901202**00
5 C PX	1.528282162**02	1.425503239**01	4.101113588**00
1 X	1.380595478**02	1.287748673**01	8.070753085**00

CUMULATIVE REGRESSIONS

NO.	VARIABLE	STD. ERROR	SUMS OF	PRGP. VAR.	F VALUE	MULTIPLE R
	NAME	OF ESTIMATE	SQUARES	= R SQ.		
2	S P/4Y	8.838378632**00	4.471646029**02	4.170922135**01	5.724297712**00	6.458267674**01
3	C P/4Y	7.33308970**00	6.956815621**02	6.48896006**01	6.468558897**00	8.055408683**01
5	C PX	6.104510893**00	8.485097782**02	7.914464145**01	7.589361498**00	8.896327413**01
1	X	4.135958695**00	9.865693260**02	9.202212017**01	1.441333997**01	9.592816488**01

PROPORTION OF VARIANCE SPECIFIED TO LIMIT VARIABLES

0

TABLE OF RESIDUALS

OBSERVATION	Y VALUE	Y ESTIMATE	RESIDUAL
1	7.970000000E+02	7.960287240E+02	9.712759852E+01
2	8.050000000E+02	8.100055308E+02	-5.005530775E+00
3	8.040000000E+02	7.995729985E+02	4.427001536E+00
4	8.060000000E+02	8.030866073E+02	2.913392693E+00
5	7.800000000E+02	7.845633745E+02	-4.563374490E+00
6	7.870000000E+02	7.870982238E+02	-9.822385013E+02
7	7.760000000E+02	7.752815191E+02	7.184808701E+01
8	7.910000000E+02	7.882794475E+02	2.720552534E+00
9	7.830000000E+02	7.845533838E+02	-1.553383812E+00
10	7.980000000E+02	7.985301906E+02	-5.301905572E+01

RATIO OF RANGES FOR THE SMALLEST RESIDUAL 5.583540324E+02

RATIO OF RANGES FOR THE LARGEST RESIDUAL 1.683587915E+01

CRITICAL VALUE OF THE RATIO AT ALPHA = .10 4.090000000E+01

BIMD31GR

Norma C. Hull

(Fortran 63)

DESCRIPTION

This program is a modification of the program "BIMD-31" written at the University of California at Los Angeles. BIMD31GR is a general linear hypothesis program with assumed model:

$$E(Y_i) = \sum_{j=1}^M X_{ij} B_j$$

$$\sigma^2(Y_i) = \frac{\sigma^2}{P_i}.$$

The maximum number of variables which can be processed is limited to 60. The upper limitation on sample size without using the graphing option for a problem is 9999; but the upper limit on sample size is 300 when graphs are specified. The sample size must be at least one greater than the number of variables. This program can perform transformations on any variable and can generate new variables if desired, according to the codes specified in the transgeneration cards. Any number of variables can be generated, but the total number of variables must not exceed the limit of 60. The program can handle balanced or unbalanced analysis-of-variance and covariance type designs. By specifying the model with the Contrast Card, as many linear contrasts can be computed as desired and the corresponding hypotheses may be tested. By using code 15 of the transgeneration list, weights may be used. There is no standard format in this program and a variable

format must be specified for the input data. The last variable must be the dependent variable. Therefore if variables are to be generated, the dependent variable must be transferred to the last variable position by using a Transgeneration Card.

The input consists of the number of variables, the number of observations or cases, and the observed values of the experiment.

Optional output includes lists of input data, tables of residuals for each contrast, and graphs based upon each contrast card where the observed value and expected value are plotted against a chosen independent variable. The output also includes for each contrast card:

- (1.) Sums of cross products matrix,
- (2.) Inverted cross products matrix,
- (3.) Regression coefficients,
- (4.) Standard deviations of regression coefficients,
- (5.) Variance of estimate,
- (6.) Standard error of estimate,
- (7.) Degrees of freedom.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ EXECUTE Card
- (3) Binary Program Deck
- (4) Δ Card
- (5) Graphing Option Card
- (6) Problem Card
- (7) Transgeneration Card(s) (if used)
- (8) Scale Card(s)
- (9) Variable Format Card(s)

- (10) Input Data Cards
- (11) Contrast Card(s)
- (12) Graph Card (if used)
- (13) Coefficient Card(s)

Card (11), (12), and (13) may be repeated for each problem, and cards (6)-(13) may be repeated for as many problems as desired.

Master Control System Card

Starting in column 1 the following characters appear:

ΔCOOP, Request Number, Name, ϕ /HD2, G.

Graphing Option Card

If graphs are desired for any fit in any problem, a "1" must be punched in column one. If no graphs are desired, a "0" (zero) must be punched in column one. This card sets up the plotting tape, and the card must precede the Problem Card for the first problem but for none of the other subsequent problems.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	PR	<u>PR</u> oblem Card
3-6		Problem number
7,8		Number of Variables ($NP \leq 60$)
9-12		Number of cases ($N\phi \leq 9999$)
13,14		Number of Transgeneration Cards

15,16	Number of Variables generated in addition to NØ above (NEW)
17,18	Number of Contrast Cards (NH)
19,20	01 if adjusting values of P are given in Input Data Cards. Otherwise, keypunch 00. (NC)
21,22	01 if Input Data are to be printed out; otherwise keypunch 00. (MD)
23,24	If input data (only input data) are to be read from a tape other than the system input tape, keypunch the tape unit number in these columns; otherwise, leave them blank.
25-70	Alphanumeric header
71,72	Number of Variable Format Cards for input data. A Number must be specified, since the program does not use a standard input format. (ND)

Transgeneration Card(s)

If a non-zero number is specified in column 13 and 14 of the Problem Card, the same number of Transgeneration Card(s) must be prepared. Different types of transformations can be performed successively on the same variable, if desired. The format of the Transgeneration card is as follows:

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	TR	TRansgeneration Card
3,4		Number of variable to be assigned to transformed or generated variable
5,6		Transformation code (codes 01, 02,...,27 of the transgeneration list, Appendix A)
7,8		Number of variable to be transgenerated or A-variable
9-14		Constant or number of B-variable (keypunch a constant with decimal point; constant may be negative.)

Scale Card(s)

The numbers punched in the Scale Card(s) are signed integers (+ need not be punched) which indicate the power of ten by which the input variable is to be multiplied. If no scaling is desired, blank fields must be indicated for each input variable. The fields are right justified, and the format of this card is I6.

<u>Columns</u>	<u>Explanation</u>
1-6	Scale factor for first variable
7-12	" " " second "
⋮	
67-72	" " " twelfth "

Scaling continues on the next card if there are more than twelve variables. Scale card(s) must be present.

Format Card(s)

Columns 1-72 of up to five cards may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats.

Input Data Cards

The sample problems will help explain the input data card preparation.

In a 2×2 factorial experiment with 3 replications, for example, the observed values may be represented as follows:

	B_1	B_2
A_1	$Y_{111} = 5$	$Y_{112} = 5$
	$Y_{211} = 8$	$Y_{212} = 5$
	$Y_{311} = 9$	$Y_{312} = 6$
	$Y_{121} = 11$	$Y_{122} = 7$
	$Y_{221} = 14$	$Y_{222} = 8$
	$Y_{321} = 15$	$Y_{322} = 6$

In general, the values can be written y_{rij} , where $i = 1, 2$; $j = 1, 2$; and $r = 1, 2, 3$.

One design parameter must be present for each analysis-of-variance variable, or cross classification population, and each design parameter is considered as a regression coefficient in terms of the model explained in the Statistical Procedure section below. In the above example, there are four analysis-of-variance variables, corresponding to the combinations A_1B_1 , A_1B_2 , A_2B_1 , and A_2B_2 ; therefore, four design parameters would be present. For this example we will denote these by μ_{11} , μ_{12} , μ_{21} , and μ_{22} .

The following shows one way of setting up the coefficients of the design parameters for the above 2×2 factorial experiment. Note that the coefficients are zero for each analysis-of-variance variable except that of the population from which the observation is taken.

Design Parameters					Dependent Variable
1	0	0	0	5	μ_{11}
1	0	0	0	8	
1	0	0	0	9	
0	1	0	0	5	μ_{12}
0	1	0	0	5	
0	1	0	0	6	
0	0	1	0	11	μ_{21}
0	0	1	0	14	
0	0	1	0	15	
0	0	0	1	7	μ_{22}
0	0	0	1	8	
0	0	0	1	6	

The first four columns are coefficients of design parameters, and the last column is the dependent variable. Each column is considered as a variable; therefore, these are five variables in the above example.

The preparation of the Input Data Cards is as follows:

- (1.) Since the program does not use the standard format, the data may be keypunched in any format.
- (2.) If there are covariates, the covariates follow the coefficients, of the design parameters in each row, and precede the dependent variable.
- (3.) Keypunch data row-wise according to the format specified. Each row may extend to the second and subsequent cards. The second row of data must start on a new card.
- (4.) Keypunch values of P following the dependent variable. In this case, 01 punch must be present in Col. 19 and 20 of the Problem Card, and a scale factor for adjusting values must be provided in the Scale Card.

Contrast Card(s)

This card, Contrast Card has a fourfold purpose:

- (1.) It indicates the number of graphs for a particular model
- (2.) It indicates the number of Coefficient Cards
- (3.) It controls output from the model fitted
- (4.) It specifies which variables are to be included in the model.

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	CO	<u>C</u> Ontrast Card
3		Number of Graphs (NGP)
4,5		Number of Coefficient Cards (NT)
6,7		Number of columns used to specify each value of coefficient of design parameter (always 6 columns on 1604)
8,9		01 if table of residuals is desired; otherwise keypunch 00. (MD)
10-68		Specify the contrast using one column for each design or covariate variable. 1 if the variable is to be included in the contrast. 0 if it is not to be included.

Different sets of variables can be made as many times as desired. This means that it is possible to perform multiple regression and correlation analyses on several linear models. The program will retain all the variables, original and generated, until all contrast cards for one problem are processed. Contrast Card(s) must be present, and the number of Contrast Card(s) prepared must agree with the number specified in columns 17, 18 of the Problem Card.

Although several Contrast Cards may be included in one problem, the arrangement of variables will not change. Therefore, the number of the variables to be included in the model must be stated in terms of the basic set of variables, which may be the original set of variables after any transgenerations are performed. If the model is not stated in terms of the basic set, the results will be meaningless.

Graph Card (if used)

If NGP is a positive number in column 3 on the Contrast Card, variable numbers must be specified to agree in number with NGP in fields of I2.

Coefficient Cards

The preparation of the Coefficient Cards is as follows:

- (1.) The coefficients are keypunched row-wise.
- (2.) Each value must occupy the same number of columns, and the number of columns must be specified in Col. 6 and 7 of the Contrast Card. (The program will generate an input format from this number.)
- (3.) Key punch the values with decimal points and minus signs, if they are negative.
- (4.) The second row (contrast) must start on a new card.

Different Contrast Cards with the associated Coefficient Cards may be repeated as many times as desired.

EXAMPLES

In the first example for BIMD31GR a 2×2 factorial experiment with three replications is shown. The first example is described in the program write.

The second example is a trigonometric regression model where

$$Y = A + Bx + C \sin(\pi/4X) + D \cos(\pi/4X) + E \sin(\pi X) + F \cos(\pi X)$$

modulo 2π . The first Contrast Card omits the "X" and " $\sin(\pi X)$ " terms from the fit. The second Contrast Card omits the " $\sin(\pi X)$ " term from

the fit. Graphs of the regression line and the data points are desired for both fits. The independent variable for both graphs is "X." Notice the intercept must be transgenerated since this program does not automatically give an intercept.

The 1604-A job card must show that tape drive "2" is used for plotting.

0
PR : 5 12 0 0 1.001 TEST CASE BIMD3IGR FOR STATPAK

(SF10.0)

1.0	0.0	0.0	0.0	5.0
1.0	0.0	0.0	0.0	8.0
1.0	0.0	0.0	0.0	9.0
0.0	1.0	0.0	0.0	5.0
0.0	1.0	0.0	0.0	5.0
0.0	1.0	0.0	0.0	6.0
0.0	0.0	1.0	0.0	11.0
0.0	0.0	1.0	0.0	14.0
0.0	0.0	1.0	0.0	15.0
0.0	0.0	0.0	1.0	7.0
0.0	0.0	0.0	1.0	8.0
0.0	0.0	0.0	1.0	6.0

CO0 6 60111111
 .5 .5 -.5 -.5
 .5 -.5 .5 -.5
 1. 1. 0. 0.
 0. 0. 1. -1.
 1. 0. -1. 0.
 0. 1. 0. -1.

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

GENERALIZED REGRESSION • TEST CASE B1MD21GR FOR STATPAK

NO. OF VARIABLES 5
 SAMPLE SIZE 12

INPUT DATA AS READ FROM CARDS

CASE NO.

1	1.0000000000E+00	0	0	0	5.0000000000E+00
2	1.0000000000E+00	0	0	0	8.0000000000E+00
3	1.0000000000E+00	0	0	0	9.0000000000E+00
4	0	1.0000000000E+00	0	0	5.0000000000E+00
5	0	1.0000000000E+00	0	0	5.0000000000E+00
6	0	1.0000000000E+00	0	0	6.0000000000E+00
7	0	0	1.0000000000E+00	0	1.1000000000E+01
8	0	0	1.0000000000E+00	0	1.4000000000E+01
9	0	0	1.0000000000E+00	0	1.5000000000E+01
10	0	0	0	1.0000000000E+00	7.0000000000E+00
11	0	0	0	1.0000000000E+00	8.0000000000E+00
12	0	0	0	1.0000000000E+00	6.0000000000E+00

CONTRAST

CARD NO. 1 1111

SUMS OF CROSS PRODUCTS

ROW 1

3.0000000000E+00 0 0 0 2.2000000000E+01

ROW 2

0 3.0000000000E+00 0 0 1.6000000000E+01

ROW 3

0 0 3.0000000000E+00 0 4.0000000000E+01

ROW 4

0 0 0 3.0000000000E+00 2.1000000000E+01

INVERTED CROSS PRODUCTS MATRIX

ROW 1

3.3333333333E+01 0 0 0

ROW 2

0 3.3333333333E+01 0 0

ROW 3

0 0 3.333333333E+01 0

ROW 4

0 0 0 3.333333333E+01

TABLE OF RESIDUALS

OBSERVATION Y VALUE Y ESTIMATE RESIDUAL

1 5.000000000E+00 7.333333333E+00 -2.333333333E+00

2 6.000000000E+00 7.333333333E+00 6.666666673E+01

3 9.000000000E+00 7.333333333E+00 1.666666667E+00

4 5.000000000E+00 5.333333333E+00 -3.333333325E+01

5 5.000000000E+00 5.333333333E+00 -3.333333325E+01

6 6.000000000E+00 5.333333333E+00 6.666666673E+01

7 1.100000000E+01 1.333333333E+01 -2.333333330E+00

8 1.400000000E+01 1.333333333E+01 6.666666697E+01

9 1.500000000E+01 1.333333333E+01 1.666666670E+00

10 7.000000000E+00 7.000000000E+00 0

11 6.000000000E+00 7.000000000E+00 1.000000000E+00

12 6.000000000E+00 7.000000000E+00 -1.000000000E+00

REGRESSION COEFFICIENTS

7.333333333E+00 5.333333333E+00 1.333333333E+01 7.000000000E+00

STANDARD DEVIATIONS OF REGRESSION COEFFICIENTS

9.128709291E+01 9.128709291E+01 9.128709291E+01 9.128709291E+01

VARIANCE OF ESTIMATE 2.500000000E+00

STD. ERROR OF ESTIMATE 1.581138830E+00

DEGREES OF FREEDOM 8

CONTRASTS AND ESTIMATES

PARAMETER

1

2

3

4

5

1	4.999999999E+01	4.999999999E+01	1.000000000E+00	0	1.000000000E+00
2	4.999999999E+01	-4.999999999E+01	-1.000000000E+00	0	0
3	-4.999999999E+01	4.999999999E+01	0	1.000000000E+00	-1.000000000E+00
4	-4.999999999E+01	-4.999999999E+01	0	-1.000000000E+00	0

ESTIMATE -3.853333333E+00 4.166666665E+00 2.000000000E+00 6.333333330E+00 -5.999999998E+00

STD, ERR, OF EST, 9,1287092913,=0| 9,12870929 3,=0| 1,2909944487,=00 1,2909944487,=00 1,2909944487,=00

PARAMETER 6

1 0

2 1,0000000000,=00

3 0

4 =1,0000000000,=00

ESTIMATE =1,6666666667,=00

STD, ERR, OF EST, 1,2909944487,=00

1
PR 3 1016 4 20001 Y=A+BX+CSINP/4*X+DC6SP/4*X+ESINP*X+FC6SP*X

TR010801 1.0

TR070803 0.0

TR0309023,1416

TR030903,25

TR0324036,2832

TR031603

TR0409023,1416

TR040904,25

TR0424046,2832

TR041704

TR0509023,1416

TR0524056,2832

TR051605

TR0609023,1416

TR0624066,2832

TR061706

(2F5.0,F10.0)

1 797

2 805

3 804

4 806

5 780

6 787

7 776

8 791

9 783

10 798

GB10000011011011

.2

GB10000011111011

2

SUMS OF CROSS PRODUCTS

ROW 1	1.0000000000000000	1.7071007335000000	7.0711954513000000	0	7.0269999999000000
ROW 2	1.7071007335000000	5.4999944899000000	4.9998889802500000	2.9288457383000000	1.4014794027000000
ROW 3	7.0711954513000000	4.9998889802500000	4.5000055699000000	-7.0711219874000000	5.0089615428000000
ROW 4	0	2.9288457383000000	-7.0711219874000000	1.0000000000000000	4.0000000000000000

INVERTED CROSS PRODUCTS MATRIX

ROW 1	1.0636306076000000	-3.1815165771000000	-1.3178697770000000	-6.4668873360000000
ROW 2	-3.1815165771000000	1.9366199096000000	-1.7305041756000000	-6.9169349474000000
ROW 3	-1.3178697770000000	-1.7605041756000000	2.2887297983000000	1.6699512115000000
ROW 4	-6.4668873360000000	-6.9169349474000000	1.6699512115000000	1.0138342922000000

TABLE OF RESIDUALS

OBSERVATION	X VALUE	Y VALUE	Y ESTIMATE	RESIDUAL
1	1.0000000000000000	7.9700000000000000	7.8952452445000000	7.4754755497000000
2	2.0000000000000000	8.0499999999000000	8.0503438613000000	-3.4386143088000000
3	3.0000000000000000	8.0399999999000000	7.9920899706000000	4.7910029292000000
4	4.0000000000000000	8.0599999999000000	8.0233934217000000	3.5606578231000000
5	5.0000000000000000	7.7999999999000000	7.8571316373000000	-5.7131637335000000
6	6.0000000000000000	7.8700000000000000	7.8594843004000000	1.7515693575000000
7	7.0000000000000000	7.7599999999000000	7.7602879024000000	-2.3790250421000000
8	8.0000000000000000	7.9099999999000000	7.8864345480000000	-2.3565451950000000
9	9.0000000000000000	7.8300000000000000	7.8952452445000000	-6.5245244511300000
10	1.0000000000000000	7.9300000000000000	8.0503438613000000	-7.7343861431000000

REGRESSION COEFFICIENTS

7.9155516010000000	9.5422274465000000	-6.8479787408000000	3.9362734451000000
--------------------	--------------------	---------------------	--------------------

STANDARD DEVIATIONS OF REGRESSION COEFFICIENTS

1.9908850758000000	2.6364141989000000	2.9204383617000000	1.9437225203000000
--------------------	--------------------	--------------------	--------------------

VARIANCE OF ESTIMATE	3.7265036940**0
STD. ERROR OF ESTIMATE	5.1045095582**01
DEGREES OF FREEDOM	6

CONTRAST	
CARD NO. 2	111101

SUMS OF CROSS PRODUCTS

ROW 1					
1.0000000000**01	5.5303000000**01	1.7071007335**00	7.0711954513**01	0	
7.9269999999**03					

ROW 2					
5.5000000000**01	3.3500000000**02	6.7771098527**00	1.0364054496**01	5.0000000000**00	
4.3452000000**04					

ROW 3					
1.7071007335**00	6.7771098527**00	5.4099944899**00	4.9998898025**01	2.9288457383**01	
1.4014794027**03					

ROW 4					
7.0711954513**01	1.0364054496**01	4.9998898025**01	4.5000055099**00	-7.0711219874**01	
5.3089615428**02					

ROW 5					
0	5.0000000000**00	2.9288457383**01	-7.0711219874**01	1.0000000000**01	
4.7000000000**01					

INVERTED CROSS PRODUCTS MATRIX

ROW 1					
5.5718950447**01	-8.1968511902**02	-8.6567357469**02	1.1900718932**01	5.1934823842**02	

ROW 2					
-8.1968511899**02	1.4933378091**02	9.5549521545**03	-2.4033817473**02	-9.4427147885**03	

ROW 3					
-8.6567357467**02	9.5549521545**03	2.0031155207**01	-3.3658818395**02	-1.3224348835**02	

ROW 4					
1.1900718932**01	-2.4033817473**02	-3.3658818395**02	2.6703092990**01	3.1927233126**02	

ROW 5					
5.1934823842**02	-9.4427147885**03	-1.3224348835**02	3.1927233126**02	1.0736629177**01	

TABLE OF RESIDUALS				
OBSERVATION	X VALUE	Y VALUE	Y ESTIMATE	RESIDUAL
1	1.0000000000**00	7.9700000000**02	7.9602872403**02	9.7127595542**01

2	2.0000000000+00	8.0497999998+02	8.1000553078+02	-5.0055307895+00
3	3.0000000000+00	8.0399999998+02	7.9957299841+02	4.4270015806+00
4	4.0000000000+00	8.0599999998+02	8.0308660732+02	2.9133926779+00
5	5.0000000000+00	7.7999999999+02	7.8456337448+02	-4.5633744895+00
6	6.0000000000+00	7.8700000000+02	7.8709822391+02	-9.8223924637+02
7	7.0000000000+00	7.7599999999+02	7.7523151920+02	7.1848079561+01
8	8.0000000000+00	7.9099999999+02	7.8827944755+02	2.7205524445+00
9	9.0000000000+00	7.8300000000+02	7.8455338388+02	-1.5533838868+00
10	1.0000000000+01	7.9800000000+02	7.9853019062+02	-5.3019063174+01

REGRESSION COEFFICIENTS

7.9944444998+02 -1.4344175194+00 8.5847850789+00 -4.5347763747+00 4.8451141015+00

STANDARD DEVIATIONS OF REGRESSION COEFFICIENTS

3.0872898730+00 5.0491503668+01 1.3510960479+00 2.1396565912+00 1.3552203724+00

VARIANCE OF ESTIMATE

1.7106134778+01

STD. ERROR OF ESTIMATE

4.1359563316+00

DEGREES OF FREEDOM

5

GENLIN
F. L. Miller, Jr.
(Fortran 63)

DESCRIPTION

This program is a modification of program "BIMD-14" written at the University of California at Los Angeles. GENLIN is designed to provide the appropriate statistics for analyzing any experimental data which may be described in terms of the general linear hypothesis model. This includes data from balanced or unbalanced experimental designs with or without proportionate subclass numbers, with or without missing observations, and with or without covariates.

Not all of the analysis of variance variables or all of the covariates need to be read in by the program. Some may be constructed or transgenerated within the program. For example, if $X_{3i} = X_{1i} - X_{2i}$, $i = 1, 2, \dots, N$, X_{3i} need not be read into the machine. It may be computed by the program provided X_{1i} and X_{2i} are stored in the memory. Similarly, if $V_{3ij} = V_{2ij}^2$, V_{3ij} may be computed by the program if V_{2ij} is stored.

Although the program will handle data from almost any linear model, it is not recommended for general use. Ordinary analysis of variance problems in balanced designs, for example, would be more easily solved by a less general program. Similarly an ordinary linear regression problem would be more satisfactorily handled by a linear regression program such as BIMD-31GR or BIMD-29GR. However, for problems which do not fall into the simpler categories GENLIN will be very useful.

An understanding of linear statistical models is necessary for the intelligent use of this program.

The General Linear Model

The general linear model may be written

$$Y_{ij} = \alpha_1 X_{1i} + \alpha_2 X_{2i} + \dots + \alpha_p X_{pi} + \beta_1 V_{1ij} + \dots + \beta_q V_{qij} + \epsilon_{ij}$$

in which Y_{ij} is the j th observation at the i th design point. The i th design point is characterized by the vector

$$(X_{1i}, X_{2i}, \dots, X_{pi})$$

in which the elements, $X_{1i}, X_{2i}, \dots, X_{pi}$, are called analysis of variance variables. The elements, $V_{1ij}, \dots, V_{qij}$, in the model are called covariates. The symbols, $\alpha_1, \alpha_2, \dots, \alpha_p, \beta_1, \dots, \beta_q$ are unknown coefficients which will be estimated from the data. The term ϵ_{ij} represents a random disturbance peculiar to the j th observation at the i th design point. These disturbances are assumed to be stochastically independent (for all i, j) and are assumed to have a common probability density function with zero mean and constant variance.

Program Restrictions

Let N represent the number of distinct design points in the problem. Then $i = 1, 2, \dots, N$. Let n_i be the number of observations at the i th design point. Then $j = 1, 2, \dots, n_i$. The program restrictions on these numbers are

$$N \leq 999$$

$$n_i \leq 999, \quad i = 1, 2, \dots, N.$$

The largest model which GENLIN will handle has 60 terms. That is

$$p + q \leq 60,$$

$$0 \leq p \leq 60,$$

$$0 \leq q \leq 60.$$

If NTR represents the number of transgenerations to be performed,

$$NTR \leq 99.$$

The program is designed to calculate the statistics for several linear models at one pass. These models are called hypotheses. The several models (or hypotheses) are all variants of the one general model described in the previous sections. They are defined by eliminating some of the terms in the full model. If H represents the number of models the user of the program wishes to consider, then the program restriction on the number of models which may be handled at one time is

$$H \leq 57.$$

Input

The input includes the number of design points, the number of analysis of variance variables, the number of covariates, the number of hypotheses, and the data punched according to the variable format card.

Output

The output from GENLIN includes

- (1) The means and standard deviations of the dependent variable (Y) and the means of the covariates (V's) for each design point.

- (2) The sum of squares explained by each hypothesis.
- (3) The residual sum of squares for each hypothesis.
- (4) The estimates of the coefficients (α 's and β 's) for each hypothesis.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ EXECUTE Card
- (3) Binary Program Deck
- (4) Δ Card
- (5) Problem Card
- (6) Scale Card(s)
- (7) Transgeneration Card(s) (if any)
- (8) Data Format Card(s)
- (9) Design Format Card(s)
- (10) Design Card(s) for 1st Design Point
- (11) Data Card(s) for 1st Design Point
- ⋮
- (N+9) Design Card(s) for Last Design Point
- (N+10) Data Card(s) for Last Design Point
- (N+11) Hypothesis Card(s).

Items 5 through N+11 may be repeated as often as desired.

Master Control System Card

Starting in column 1 the following characters appear:

Δ COOP, REQUEST NUMBER, NAME.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	PR	Problem
3-5		Number of design points, N.
6,7		Number of analysis of variance variables, p.
8,9		Number of covariates, q.
10,11		Number of hypothesis cards, H.
12,13		Number of transgeneration cards, NTR.
14	0	Test of accuracy of coefficients not desired.
	1	Test of accuracy of coefficients desired.
15		Number of data format cards, 1,2,3,4 or 5.
16		Number of design format cards, 1,2,3,4, or 5.
17-72		Any alphanumeric character(s) to form a title.

The numbers p and q are the number of analysis of variance variables and the number of covariates in the largest model even if some of them are to be constructed by transgeneration.

Scale Card

<u>Columns</u>	<u>Explanation</u>
1-4	Scale factor for first covariate or if $q = 0$, the scale factor for the dependent variable.
5-8	Scale factor for the second covariate or if $q = 1$, the scale factor for the dependent variable.
9-12	
⋮	
69-72	and so forth

The numbers punched in the Scale Card are signed integers (+ need not be punched) which indicate the power of ten by which the covariate or dependent variable is to be multiplied. If the field is blank, no scaling will be performed.

There must be at least one Scale Card. If $q \leq 18$, more than one Scale Card will be required, the additional cards being interpreted as a continuation of the first Scale Card.

The program considers the list of covariates and dependent variable to be arranged in order with the dependent variable last. The scale factors on the Scale Cards should follow the same order.

Transgeneration Cards

It is important that the number of Transgeneration Cards used agree with the number punched in columns 12 and 13 of the Problem Card.

For the purpose of transgeneration the program considers the analysis of variance variables and the covariates and independent variable to be numbered sequentially. Thus the analysis of variance variables are numbered 1, 2, ..., p. The covariates are numbered p+1, p+2, ..., p+q, and the dependent variable is numbered p+q+1. The transgenerations which can be used in GENLIN are listed in Appendix B.

Data Format Card

The Data Format Card describes the format in which the covariates and dependent variable are to be read in. Only F-type FORTRAN formats may be used. See Appendix C for details. There are two cases to consider.

No Covariates

If $q = 0$, the n_i observations of the dependent variable at the i th design point may be grouped sequentially on one card or more. For example, if there were at most 10 values of the dependent variable at a design point the format

(10F6.0)

punched in columns 1 through 8 of the Data Format Card would be appropriate. If there were 14 values of the dependent variable at a design point with 12 values on one card and 2 on the following card, an appropriate format would be

(12F6.0/2F6.0)

punched in columns 1 through 14 of the Data Format Card.

It is important that the format chosen be such that it will accommodate the largest of the n_i , $i = 1, 2, \dots, N$.

With Covariates

If $q \geq 1$, there must be n_i (or if more than one card per observation is needed, a multiple of n_i) data cards for each design point. The covariates are considered to be listed first and the dependent variable last on each of these. The Data Format Card must provide for reading in $q+1$ values even though some of the covariates will be constructed within the program by transgeneration.

Design Format Card

The Design Format Card must allow for the number of observations at the design point (n_1) in the first field and for p analysis of variance variables in the subsequent fields, whether some of the analysis of variance variables are to be transgenerated or not.

The printed output will include an N by p table of the analysis of variance variables. Each entry in this printed table will consist of a signed 3-digit integer.

Design Cards

The number of Design Cards must be N , or a multiple of N if it is not possible to punch all of the analysis of variance variables for a design point on one card. The format for the Design Card must follow the format specified on the Design Format Card.

As stated in the previous section, the first number punched in the Design Card is the number of observations (n_1) for the design point. Then there follows p fields for p analysis of variance parameters. If one or more of the analysis of variance parameters are to be constructed by transgeneration, the corresponding field in the Design Card may be left blank.

Data Cards

The Data Cards are punched according to the format on the Data Format Card. There are two cases:

No Covariates

As before, let Y_{ij} represent the j th observation at the i th design point. Then the sequence of observations punched on the Data Card is

$$Y_{i1} \ Y_{i2} \ \dots \ Y_{in_i}$$

The format is as specified on the Data Format Card. The number of observations punched on the Data Card for the i th design point must agree with the first number punched on the Design Card for the i th design point.

With Covariates

Let $V_{1ij}, V_{2ij}, \dots, V_{qij}$ represent the j th set of q covariates at the i th design point and let Y_{ij} represent the j th observation at the i th design point. Then there will be a set of Data Cards for the i th design point punched as follows

$$\begin{array}{ccccccccc} V_{1i1} & V_{2i1} & \dots & V_{qi1} & Y_{i1} & & & & \\ V_{1i2} & V_{2i2} & \dots & V_{qi2} & Y_{i2} & & & & \\ . & . & . & . & . & . & . & . & . \\ V_{1in_i} & V_{2in_i} & \dots & V_{qin_i} & Y_{in_i} & & & & \end{array}$$

(A single row in the above may represent one or more Data Cards depending on the format specified in the Design Format Card.)

The number of rows in the preceding table must agree with the first number punched in the Design Card for the i th design point. The format for these cards is that format specified on the Data Format Card.

Hypothesis Cards

As many as 57 Hypothesis Cards may be placed at the end of the Design and Data Cards. A Hypothesis Card defines a specific linear model by designating which of the $p+q$ terms in the full model are to be included and which are to be excluded from consideration. The Hypothesis Cards are punched as follows:

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1	H	Hypothesis Card
2	0	if 1st term is excluded
	1	if 1st term is included
3	0	if 2nd term is excluded
	1	if 2nd term is included
.....		
$p+q+1$	0	if last term is excluded
	1	if last term is included
62-72		any alphanumeric identification.

Three hypotheses or models are programmed and hence Hypothesis Cards for these models need not be included in the set of Hypothesis Cards. These are the models containing

- a. no terms whatsoever
- b. $p+q$ terms (the full model)
- c. The first term only.

EXAMPLES

Two examples of the use of GENLIN are provided on the following pages. The first is an example of the use of GENLIN when there are

no covariates and the second is an example with covariates. Both are simple examples and could be worked on other programs. Simple examples were chosen because they would be easy to follow.

Example 1

The first example is an illustration of the analysis of a 2^4 factorial experiment with three observations per cell. The experimental design is given in the following table

	Factor				<u>Observations</u>
	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	
(1)	-	-	-	-	6.07, 6.08, 6.09
a	+	-	-	-	6.04, 6.03, 6.05
b	-	+	-	-	
ab	+	+	-	-	
c	-	-	+	-	
ac	+	-	+	-	
bc	-	+	+	-	
abc	+	+	+	-	
d	-	-	-	+	
ad	+	-	-	+	
bd	-	+	-	+	
abd	+	+	-	+	
cd	-	-	+	+	
acd	+	-	+	+	
bcd	-	+	+	+	
abcd	+	+	+	+	6.24, 6.23, 6.22

In the table "-" means that the factor is controlled at one level and "+" means that the factor is controlled at the other level. On the Design Cards the "-" is represented by -1 and the "+" by 1. A is in

columns 7 and 8, B is in columns 11 and 12, C is in columns 19 and 20, and D is in columns 35 and 36. The interactions among the factors have been coded into the analysis by transgeneration.

The sum of squares for an analysis of variance table may be obtained from the printed output of GENLIN.

Analysis of Variance

<u>Source</u>	<u>d.f.</u>	<u>Sum of Squares</u>	<u>Hypothesis Number</u>
Total	48	1,958.1005	1 residual
Mean	1	1,954.9598	20 explained
A	1	.4389	3 "
B	1	.0054	4 "
C	1	.0698	5 "
D	1	.8993	6 "
2-Factor Int.	6	1.1442	13 "
AB	1	.00001	7 "
AC	1	.0137	8 "
BC	1	.0527	9 "
AD	1	.3120	10 "
BD	1	.7575	11 "
CD	1	.0083	12 "
3-Factor Int.	4	.5728	18 "
ABC	1	.2655	14 "
ABD	1	.1230	15 "
ACD	1	.0005	16 "
BCD	1	.1838	17 "
ABCD	1	.0042	19 "
Residual	32	.0062	2 residual

The total sum of squares (1,958.1005) with 48 degrees of freedom is read from the output as the residual sum of squares for Hypothesis 1.

The sum of squares for the mean (1,954.9598) is read from the output as the sum of squares explained by Hypothesis 20. The other sums of squares are obtained in a similar manner.

Example 2

The second example is an illustration of the use of GENLIN for analyzing a 5×5 Latin Square with a covariate. The data were taken from "Statistical Methods," George W. Snedecor, The Iowa State College Press, page 337.

There are 13 analysis of variance variables. The first measures the general mean. The next four measure the effects of the differences between Row 1 and Row 5, between Row 2 and Row 5, and so forth. The next four analysis of variance variables measure the differences between Column 1 and Column 5, Column 2 and Column 5, and so forth. The last four measure the differences between Treatment 1 and Treatment 5, Treatment 2 and Treatment 5 and so forth.

The Hypothesis Cards were constructed to provide an analysis of variance ignoring the covariate as well as an analysis with the covariate.

In the analysis of variance tables which follow, in the last column, E represents sum of squares explained and R represents residual sum of squares.

Analysis of Variance Ignoring Covariate

<u>Source</u>	<u>d.f.</u>	<u>Sums of Squares</u>	<u>Hypothesis Number</u>	
Total	25	1,626,188.	1	R
Mean	1	1,589,616.64	13	E
Rows	4	13,601.36	4	E
Columns	4	6,146.16	6	E
Treatments	4	4,156.56	8	E
Residual	12	12,667.28	3	R

The total of the sums of squares for rows, columns and treatments in the above analysis is the sum of squares explained by Hypothesis 12.

Analysis of Variance With Covariate

<u>Source</u>	<u>d.f.</u>	<u>Sums of Squares</u>	<u>Hypothesis Number</u>
Total	25	1,626,188	1 R
Mean	1	1,589,616.64	13 E
Covariate	1	215.52	10 E minus 13 E
Rows*	4	8,554.59	5 R " 2 R
Columns*	4	6,152.36	7 R " 2 R
Treatments*	4	3,864.80	9 R " 2 R
Residual*	11	12,185.56	2 R

*These effects are adjusted for the covariate and are not orthogonal.

4 1 1 1 1 1 1 1 1 1 1
H
4 1 1 1 1 1 1 1 1 1 1
H 1 4 1 1 1 1 1 1 1

ALL BUT TREATMENTS
MEAN, COVARIATE
TOTALING MEAN
ROW, COL, TRT

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

GENERAL LINEAR HYPOTHESIS FOR ANOVA

PROBLEM "LATIN SQUARE WITH COVARIATE, EXAMPLE FROM SNEDECOR, P337

NUMBER OF DESIGN CARDS 25

NUMBER OF PARAMETERS 14

DESIGN PATTERNS

1	1	1	-0	-0	-0	1	-0	-0	-0	-0	1	-1	-0
2	1	1	-0	-0	-0	-0	10	-0	-0	-1	-1	-1	-1
3	1	1	-0	-0	-0	-0	1	-0	-0	1	-0	-1	-0
4	1	1	-0	-0	-0	-0	-0	10	-0	-0	-0	1	-0
5	1	1	-0	-0	-0	-1	-1	-1	-1	-0	-0	-1	1
6	1	-0	1	-0	-0	1	-0	-0	-0	-0	-0	-1	1
7	1	-0	1	-0	-0	-0	10	-0	-0	1	-0	-1	-0
8	1	-0	1	-0	-0	-0	1	-0	-0	-1	-1	-1	-1
9	1	-0	1	-0	-0	-0	-0	-0	1	-0	1	-1	-0
10	1	-0	1	-0	-0	-1	-1	-1	-1	-0	-0	1	-0
11	1	-0	-0	1	-0	1	-0	-0	-0	-1	-1	-1	-1
12	1	-0	-0	1	-0	-0	1	-0	-0	-0	1	-1	-0
13	1	-0	-0	1	-0	-0	-0	1	-0	-0	-0	1	-0
14	1	-0	-0	1	-0	-0	-0	-0	1	-0	-0	-0	1
15	1	-0	-0	1	-0	-1	-1	-1	-1	1	-0	-0	-0
16	1	-0	-0	-0	1	1	-0	-0	-0	1	-0	-0	-0
17	1	-0	-0	-0	1	-0	1	-0	-0	-0	-0	1	-0
18	1	-0	-0	-0	1	-0	-0	1	-0	-0	-0	-0	1
19	1	-0	-0	-0	1	-0	-0	-0	1	-1	-1	-1	-1
20	1	-0	-0	-0	1	-1	-1	-1	-1	-0	1	-0	-0
21	1	-1	-1	-1	-1	1	-0	-0	-0	-0	-0	1	-0
22	1	-1	-1	-1	-1	-0	1	-0	-0	-0	-0	-0	1
23	1	-1	-1	-1	-1	-0	-0	1	-0	-0	1	-0	-0
24	1	-1	-1	-1	-1	-0	-0	-0	1	1	-0	-0	-0

HYPOTHESIS	SUM OF SQUARES		
CARD	EXPLAINED	HYPOTHESIS	
1	0	PROGRAMMED	00000000000000
2	1.6145725382**0.6	PROGRAMMED	11111111111111
3	1.6141893469**0.5	TOTAL, NO CO	11111111111111
4	1.3601360000**0.4	ROWS, NO COV	01111000000000
5	1.6056918947**0.5	ALL BUT ROW	10000111111111
6	1.9753941160**0.5	COLS, NO COV	00000111100000
7	1.6078500829**0.5	ALL BUT COL	11111000011111
8	4.1565599999**0.3	TRYS, NO COV	00000000011110
9	1.6106085044**0.5	ALL BUT TRE	11111111000001
10	1.5898321555**0.5	MEAN, COVARI	10000100000000
11	1.6019014406**0.5	TOTAL, NO ME	01111111111111
12	2.8820520583**0.5	ROW, COL, TRY	01111111111110
13	1.5896166400**0.5	PROGRAMMED	10000100000000

ESTIMATES OF COEFFICIENTS, RESIDUAL SUMS OF SQUARES, AND F TESTS

HYPOTHESIS	1	2	3	4	5
VARIABLE					
1 *TAPE PARITY ERROR NEXT LINE(S)*	0	2.1380269973E+02	2.5097024469E+02	0	1.4032049285E+02
2 *TAPE PARITY ERROR NEXT LINE(S)*	0	6.65890908689E+00	7.2365620043E+00	4.4888888886E+00	0
60 T 63 *TAPE PARITY ERROR NEXT LINE(S)*					

PE T 63	3	0	9<1884464689#*00	1<2947296112#*01	10440000000#*01	0
TAPE PARITY ERROR NEXT LINE(S)						
PE T 63	4	0	452824414023#*00	8<8602446024#*00	1<014888888#*01	0
	5	0	3.0525410273#*01	3.370244702#*01	3.496000000#*01	0
	6	0	2.6882358196#*01	2.6531567519#*01	0	2.7996971673#*01
	7	0	1.9824474724#*01	1.3227314584#*01	0	4.5156793380#*01
	8	0	4.9845199524#*00	5.2649726153#*00	0	3.4079491537#*00
	9	0	1.1807950775#*01	9.7304355693#*00	0	1.8208119870#*01
	10	0	2.7535451793#*01	2.0147296112#*01	0	4.2158881477#*01
	11	0	1.2604196951#*01	1.1829755297#*01	0	1.3410030785#*01
	12	0	6.6685759269#*00	7.1541028193#*00	0	3.9563942172#*00
	13	0	1.3140619212#*01	1.270244702#*01	0	1.4528962678#*01
	14	0	1.0042247053#*01	0	0	3.0096433875#*01
RESIDUAL SS	1.6261880000#*06	1.1610461822#*04	1.1998653107#*04	1.6125866400#*06	2.0496105285#*04	
RESIDUAL DF	25	11	12	21	15	
F TESTS			3.6778073011#*01	1.5167990928#*02	2.1046121935#*00	
F TESTS DF			1	10	4	11
HYPOTHESIS	6	7	8	9	10	
VARIABLE						
1	0	2.1360521494#*02	0	2.8480450491#*02	2.3685604369#*02	
2	0	1.9954232946#*00	0	5.5380970970#*00	0	
3	0	7.1818491500#*00	0	1.5305739031#*01	0	
4	0	6.4214807553#*00	0	1.2364787585#*01	0	
5	0	3.1680963568#*01	0	3.5763063660#*01	0	
6	2.1702048476#*01	0	0	2.6720932568#*01	0	
7	2.5143681744#*01	0	0	2.1274846134#*02	0	
8	2.3460725802#*01	0	0	4.5540858638#*00	0	
9	4.3386310595#*01	0	0	9.4125814561#*00	0	
10	0	2.5973347346#*01	1.7640000000#*01	0	0	
11	0	1.1350109800#*01	1.0640000000#*01	0	0	
12	0	3.6568188816#*01	2.4030000004#*01	0	0	
13	0	1.4440369314#*01	1.3960000000#*01	0	0	
14	0	1.0442791116#*01	0	9.1300891805#*00	4.1451669295#*00	
RESIDUAL SS	1.4286485884#*06	1.8337917013#*04	1.6220314400#*05	1.5579495575#*04	3.6355844482#*04	
RESIDUAL DF	21	15	21	15	23	
F TESTS	1.3425322464#*02	1.593433761#*00	1.5257472984#*02	9.4008687905#*01	1.9536920912#*00	
F TESTS DF	10	4	10	4	12	11
HYPOTHESIS	11	12	13			
VARIABLE						
1	0	0	2.5215999999#*02			
2	1.1383282572#*01	1.07520733 9#*02	0			
3	1.2357298696#*01	5.6674554459#*00	0			
4	1.6089189491#*01	3.0662732879#*01	0			
5	1.1660938718#*01	5.86273287#*00	0			

6	-2.8757405399	+01	-1.6171075876	+01	0
7	2.4375331926	+00	3.2221345930	+01	0
8	3.7549247779	+00	3.7562467585	+01	0
9	2.2643221139	+01	7.5641661861	+01	0
10	6.9266390152	+01	1.5325345537	+00	0
11	1.2445162684	+01	5.1462732878	+01	0
12	-4.3841370791	+00	-4.9190534364	+01	0
13	-1.4798351314	+01	2.6862732877	+01	0
14	6.7048368326	+01	0		0
RESIDUAL SS	2.4286559326	+04	1.3379827942	+06	3.6571360016
RESIDUAL DF		12		13	24
F TESTS	1.2009606149	+01	6.2831676631	+02	1.8191145479
P TESTS DF		1	2	1	1

ACD
BCD
3-FACTOR
ABCD

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

GENERAL LINEAR HYPOTHESIS FOR ANOVA

PROBLEM EXAMPLE OF GENLIN ANALYZING A FACTORIAL EXP. NO CHVAR.

NUMBER OF DESIGN CARDS 16

TRANSGENERATIONS

4 =	2 TIMES	3
6 =	2 TIMES	5
7 =	3 TIMES	5
8 =	4 TIMES	5
10 =	2 TIMES	9
11 =	3 TIMES	9
12 =	4 TIMES	9
13 =	5 TIMES	9
14 =	6 TIMES	9
15 =	7 TIMES	9
16 =	8 TIMES	9

NUMBER OF PARAMETERS 16

DESIGN PATTERNS

1															
2															
3															
4															
5															
6															
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															

DESIGN NO.	OF	MEAN	STD. DEV.
POINT	OBS.	OF Y	OF Y
1	3	6.000000001	1.0000022550

2	3	6.0400000000**00	1.0000094638**02
3	3	6.5299999999**00	2.0000045100**02
4	3	6.4299999999**00	1.0000094638**02
5	3	6.3100000000**00	0
6	3	6.0900000000**00	2.0000076144**02
7	3	6.1200000000**00	1.0000022550**02
8	3	6.3400000000**00	1.0000146726**02
9	3	6.7900000000**00	9.9999604619**03
10	3	6.6800000000**00	0
11	3	6.7300000000**00	3.0000005562**02
12	3	6.0800000000**00	2.0000014056**02
13	3	6.1700000000**00	9.9999604619**03
14	3	6.3800000000**00	1.0000022550**02
15	3	6.4899999999**00	1.0000094638**02
16	3	6.2300000000**00	9.9999604619**03

HYPOTHESIS	SUM OF SQUARES	
CARD	EXPLAINED	HYPOTHESIS
1	0	PROGRAMMED
2	1.9580943000**03	PROGRAMMED
3	4.3891874948**01	
4	5.4187500002**03	
5	6.9768749810**02	
6	8.9926874926**01	
7	1.8750000616**05	
8	1.3668749994**02	
9	5.2668750065**02	
10	3.1201874990**01	
11	7.5751875011**01	
12	8.2687501006**03	
13	1.1441625002**00	
14	2.6551874979**01	
15	1.2301875004**01	
16	4.6875000377**04	
17	1.8376875002**01	
18	5.7277499996**01	
19	4.2187500253**03	
20	1.9549597688**03	PROGRAMMED

ESTIMATES OF COEFFICIENTS, RESIDUAL SUMS OF SQUARES, AND F TESTS

HYPOTHESIS	1	2	3	4	5
VARIABLE					
1	0	6.3818750001**00	0	0	0
2	0	9.5624999942**02	9.5624999942**02	0	0
3	0	1.0625000000**02	0	1.0625000000**02	0
4	0	6.2500001027**04	0	0	0
5	0	3.8124999948**02	0	0	3.8124999948**02
6	0	1.6874999996**02	0	0	0
7	0	3.3125000020**02	0	0	0
8	0	7.4374999971**02	0	0	0
9	0	1.3687499994**01	0	0	0
10	0	8.0624999996**02	0	0	0
11	0	1.2562500001**01	0	0	0
12	0	5.0625000007**02	0	0	0
13	0	1.3125000080**02	0	0	0

14	0	3.1250000125**03	0	0	0	
15	0	6.1875000003**02	0	0	0	
16	0	9.3750000279**03	0	0	0	
RESIDUAL SS	1.9581005000**03	6.1999857425**03	1.9576615813**03	1.9580950813**03	1.9580307313**03	
RESIDUAL DF	48	32	47	47	47	
F TESTS			6.7360340063**05	6.7375256225**05	6.7373042025**05	
F TESTS DF			15	32	15	32

HYPOTHESIS VARIABLE	6	7	9	9	10	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	$\approx 6.2500001027 \approx 04$	0	0	0	
5	0	0	0	0	0	
6	0	0	$1.6874999996 \approx 02$	0	0	
7	0	0	0	$-3.3125000020 \approx -02$	0	
8	0	0	0	0	0	
9	$1.3687499994 \approx 01$	0	0	0	0	
10	0	0	0	0	$\approx 8.0624999986 \approx 02$	
11	0	0	0	0	0	
12	0	0	0	0	0	
13	0	0	0	0	0	
14	0	0	0	0	0	
15	0	0	0	0	0	
16	0	0	0	0	0	
RESIDUAL SS	$1.9572012313 \approx +03$	$1.9581004813 \approx +03$	$1.9580868313 \approx +03$	$1.9580478313 \approx +03$	$1.9577884813 \approx +03$	
RESIDUAL DF	47	47	47	47	47	
F TESTS	$6.7344500027 \approx +05$	$6.7375442034 \approx +05$	$6.7374972353 \approx +05$	$6.7373630416 \approx +05$	$6.7364706524 \approx +05$	
F TESTS DF	15	32	15	32	15	32

HYPOTHESIS	11	12	13	14	15
VARIABLE					
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	6.2500001027**04	0	0
5	0	0	0	0	0
6	0	0	1.6874999996**02	0	0
7	0	0	3.3125000020**02	0	0
8	0	0	0	7.4374999971**02	0
9	0	0	0	0	0
10	0	0	0	0	0
11	1.2562500001**04	0	1.2562500001**04	0	0
12	0	0	0	0	5.0625000007**02
13	0	1.3125000080**02	1.3125000080**02	0	0
14	0	0	0	0	0
15	0	0	0	0	0

	16	0	0	0	0	0
RESIDUAL SS	1.9573429813e+03	1.9580922313e+03	1.9589563375e+03	1.9578349813e+03	1.9579774813e+03	
RESIDUAL DF	47	47	42	47	47	
F TESTS	6.7349377456e+05	6.7375158160e+05	1.0100411033e+05	6.7366506527e+05	6.7371209766e+05	
F TESTS DF	15	32	10	32	15	32
HYPOTHESIS	16	17	18	19	20	
VARIABLE						
1	0	0	0	0	6.3818750001e+00	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	
5	0	0	0	0	0	
6	0	0	0	0	0	
7	0	0	0	0	0	
8	0	0	7.4374999971e+02	0	0	
9	0	0	0	0	0	
10	0	0	0	0	0	
11	0	0	0	0	0	
12	0	0	5.0625000007e+02	0	0	
13	0	0	0	0	0	
14	-3.1250000125e-03	0	-3.1250000125e-03	0	0	
15	0	6.1875000003e+02	6.1875000003e+02	0	0	
16	0	0	0	9.3750000279e+03	0	
RESIDUAL SS	1.9581000313e+03	1.9579167313e+03	1.9575277250e+03	1.9580362813e+03	3.1407312155e+00	
RESIDUAL DF	47	47	44	47	47	
F TESTS	6.7375426546e+05	6.7369119436e+05	8.4194667805e+05	6.7375297515e+05	1.0785508603e+03	
F TESTS DF	15	32	12	32	15	32

ACCURACY OF COEFFICIENTS						
ROW						
1	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
2	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
3	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
4	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
5	0	0	0	0	0	0

	0	0	0	0	0	0
	0	0	0	0	0	0
6	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
7	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
8	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
9	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
10	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
11	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
12	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
13	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
14	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
15	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
16	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0

NONLIN

Norma C. Hull
(Fortran 63)

DESCRIPTION

This program is a modification of A Generalized Least Squares Program For the IBM 7090 Computer by M. H. Lietzke as described in the ORNL Report 3259. The program uses a Gauss-Newton iterative search method for finding parameters (coefficients). The program will attempt to perform a least squares fit to any function, linear or non-linear, with up to eight parameters and up to five independent variables. A transgeneration subroutine package will allow the user to generate any or all of the independent or dependent variables. Weighting factors may either be defined or computed. The user must furnish a subroutine to define the function to be fitted and to define the partial derivatives of the function with respect to the coefficients.

NONLIN will plot the data and the estimated regression line. The user has the option of choosing which independent variable is to be measured along the X axis. As many as five different choices of the independent variable may be made for the same set of data. The format for the input data is variable, which permits flexibility in the lengths of the data fields. The input includes the number of data points, the number of parameters, the number of independent variables, the number of iterations allowed on a given case, Epsilon (maximum allowed percentage difference between successive values of the coefficients at convergence), the number of transgenerations to be performed, the number of graphs

desired for each problem, the variable numbers to be used as independent variables, and the observations from the experiment.

The output includes a list of the transgenerations (if any), the variance of fit and the parameters at each iteration, the final estimated parameter and its standard error, and the observed dependent variable and the calculated dependent variable, the variance of the calculated variable, and the independent variable in tabular form for each point. The user has the option of having the inverse matrix printed out.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ FTN, L, E, G.
- (3) SUBRT Fortran Subroutine
- (4) Δ EXECUTE Card
- (5) Binary Program Deck
- (6) Δ Card
- (7) Problem Card
- (8) Parameter or Coefficient Card
- (9) Variable Name Card
- (10) Format Card(s)
- (11) Input Data Cards
- (12) Transgeneration Cards (if any)

Items 7-12 may be repeated for as many problems as desired.

Master Control System Card

Starting in Column 1 the following characters appear:

ΔCOOP, REQUEST NUMBER, NAME, ϕ /HD2, G.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1-48		Alphanumeric Problem Identification (TTTL)
49-52		Number of data points $NP \leq NDP \leq 250$
53-54		Number of parameters or coefficients $1 \leq NP \leq 8$
55-56		Number of independent variables $1 \leq NX \leq 5$
57-60		Number of iterations on a given case (NIT)
61-62		Number of transgenerations (NTR)
63-70		Epsilon, the maximum allowed fractional difference between successive values of the coefficients at convergence (EPS)
71-72	0	Do not list Inverse Matrix
	1	List Inverse Matrix (NINV)
73		Number of variable Format Cards (NFMT)
74		Number of graphs for this problem (NG)
75-79		The variables which are to be considered as the independent variables for the graphs ($1 \leq ICX(I) \leq 5$)*

*The number of columns used in this field must correspond to the number of graphs which are to be plotted.

Parameter Card

<u>Columns</u>	<u>Explanation</u>
1-10	Initial estimate on first parameter
11-20	" " " second "
21-30	" " " third "
⋮	
71-80	" " " eighth "

Variable Name Card

The purpose of the Variable Name Card is to identify the variables in the output by their names. The independent variable names must be listed first in the order that the variables appear after transgeneration and then the dependent variable name must be listed. The format of the Variable Name Card is as follows:

<u>Columns</u>	<u>Explanation</u>
1-8	Name of 1st independent variable
9-16	" " 2nd " " or dependent variable
17-24	" " 3rd " " " " "
25-32	" " 4th " " " " "
33-40	" " 5th " " " " "
41-48	Name of dependent variable.

Format Card(s)

Columns 1-80 of as many as five cards may be used to describe the input data using E or F conversions. See Appendix C for details.

concerning formats. A format must be specified, and the data must be arranged in the order outlined in the description of the Data Cards.

Data Cards

The format for this program is variable; and therefore, the data must be punched in the same format as specified on the format card(s). The weighting factor, the independent variables, and the dependent variable must be punched in this order. If the weighting factor is defined in the subroutine SUBRT, the weighting factor field may be left blank; but a field for the weighting factor must be defined in the format card(s).

Transgeneration Cards (if any)

Transgeneration Cards are explained in Appendix B. It is necessary that the number of Transgeneration Cards used agrees with the number punched in Columns 61-62 of the Problem Card.

SUBRT Subroutine

In this subroutine YC, DERIV(I), and WI need to be defined. YC is the value of the dependent variable computed using the current estimates of the coefficients. This is the equation of the model for which the regression coefficients are to be computed. GP1(I) are the initial estimates of the coefficients, X(J,I) is the array of independent variables where the J subscript is the index on the data point, and the I subscript is the index of the independent variable.

DERIV(I) (I = 1,2,...,NP) is the partial derivative of the function with respect to the coefficient. If weights are read in for each data point, then the statement $WI = W(J)$ would be included in the subroutine. If weights are unity, then $WI = 1.0$ is the correct statement. In like manner if the weight is to be computed, then the calculation must be inserted. For example, if the square root of the dependent variable is desired as the weight, the statement $WI = \text{SQRTF}(Y(J))$ is correct. SUBRT must be written each time a new model is to be run. However, if a satisfactory model has been defined, then as many sets of data as desired may be run without any changes in SUBRT.

EXAMPLE

In the example for NONLIN the non-linear estimation requires that the model

$$\log T = A + B \log \sigma + C(1/\sigma + D)$$

be used. $\log \sigma$, $\log T$ and the reciprocal of σ are generated by the use of Transgeneration Cards. The restriction that D be always positive is provided for in the subroutine SUBRT. Three graphs of the fitted $\log T$ and the data points with the independent variables being chosen as σ , $\log \sigma$, and the reciprocal of σ are desired.

The 1604A job card must show that tape drive "2" is used for plotting.

```

1  SUBROUTINE SHIRT(J,W,X,Y,GPI,DERIV,YC,FI,WI)
    DIMENSIONW(250),X(250,5),Y(250),GPI(8),DERIV(8)
    IF(GPI(4) .LT. 0.0)10,20
10  GPI(4)=.00001
20  YC=GPI(1)+GPI(2)*X(J,2)+GPI(3)*(1.0/(X(J,1)+GPI(4)))
    FI=Y(J)+YC
    DERIV(1)=1.0
    DERIV(2)=X(J,2)
    DERIV(3)=1.0/(X(J,1)+GPI(4))
    DERIV(4)=GPI(3)*(X(J,1)+GPI(4))**(-2)
    WI=1.0
    RETURN
    END

```

CREEP RUPTURE PROBLEM $\ln(T) = A + B \ln(S) + C(1/S + D)$ 13 4 3 50 3.0001 113123
 44.0 -9.5 -144.5 1.0
 SIGMA LN SIGMA1/SIGMA LN T
 (5F10.0)

75.0	3.3
70.0	4.5
67.0	7.7
60.0	23.0
57.4	23.3
55.0	39.5
52.0	52.3
40.0	312.3
39.0	243.6
30.0	913.7
30.0	1032.5
25.0	2436.5
20.0	5596.7

TR 5= LN 5
 TR 3= LN 2
 TR 4= RECIP 2

TRANSGENERATIONS

5	=	LN	5
3	=	LN	2
4	=	RECIP	2

VARIANCE OF FIT

PARAMETERS

	1	2	3	4
1	5.71523E+02	4.40000E+01	9.50000E+00	1.44500E+02
2	3.04274E+01	7.81292E+01	1.60067E+01	6.60223E+02
3	3.71451E+02	1.8261E+02	2.30199E+01	1.71264E+03
4	8.99439E+00	8.21431E+01	1.67269E+01	7.49454E+02
5	9.17894E+02	9.05396E+01	1.82175E+01	9.49295E+02
6	2.26699E+00	1.12383E+02	2.20043E+01	1.55419E+03
7	2.69927E+02	1.22970E+02	2.37859E+01	1.90630E+03
8	2.00356E+02	1.22999E+02	2.37904E+01	1.90755E+03

CREEP RUPTURE PROBLEM $LN(T)=A+BLN(S)+C(1/S+D)$

PARAMETER		STD ERROR				
A1	12.30027E+01	43.00372E+00				
A2	23.79105E+00	76.30050E+01				
A3	19.07677E+02	11.02246E+02				
A4	24.27806E+00	51.00778E+01				
VARF	20.03556E+03	VARF1	15.41197E-04			
Y(OBS)	Y(CALC)	VAR(Y)	X(1)	X(2)	X(3)	
11.93922E-01	10.69658E-01	11.24615E-03	75.00000E+00	43.17488E-01	13.33333E+03	
15.04077E-01	16.91984E-01	52.71648E-04	70.00000E+00	42.48495E-01	14.28571E+03	
20.41220E-01	20.69052E-01	37.15855E-04	67.00000E+00	42.04693E-01	14.92537E+03	
31.35494E-01	29.58459E-01	33.89993E-04	60.00000E+00	40.94345E-01	15.66667E+03	
31.48453E-01	32.91868E-01	37.69850E-04	57.40000E+00	40.50044E-01	17.42160E+03	
36.76301E-01	36.00948E-01	41.18154E-04	55.00000E+00	40.07333E-01	18.19182E+03	
39.56995E-01	39.88977E-01	44.15020E-04	52.00000E+00	39.51244E-01	19.23077E+03	
57.43964E-01	55.61899E-01	40.72979E-04	40.00000E+00	36.88879E-01	25.00000E+03	
54.95528E-01	56.15162E-01	40.81390E-04	39.60000E+00	36.78829E-01	25.25253E+03	
68.17502E-01	69.38290E-01	58.10073E-04	30.00000E+00	34.01197E-01	33.33333E+03	
69.39738E-01	69.38290E-01	58.10073E-04	30.00000E+00	34.01197E-01	33.33333E+03	
77.98318E-01	77.09782E-01	60.65763E-04	25.00000E+00	32.18876E-01	40.00000E+03	
86.29932E-01	86.47078E-01	18.37351E-03	20.00000E+00	29.95732E-01	50.00000E+03	

INVERSE MATRIX

92.30190E+03	16.37450E+03	-23.59788E+05	10.45503E+03			
16.37450E+03	29.05718E+02	41.81204E+04	18.45086E+02			
-23.59788E+05	41.81204E+04	60.63955E+06	-27.32148E+04			
10.45503E+03	18.45086E+02	-27.32149E+04	12.98588E+02

ANOVA

F. L. Miller, Jr.

(FORTRAN 63)

DESCRIPTION

The program, ANOVA, performs a weighted analysis of variance for a one classification variable. The program handles unequal treatment sample sizes. Transgeneration of the data is allowed; and several problems may be run, one behind another.

The input includes the number of input preparatory variables, the number of observations, the number of treatments, and the number of transformations to be performed. The output includes an analysis of variance table; chi square and corrected chi square statistics for Bartlett's test of homogeneity of variance, the number of observations per treatment, the mean, variance of the mean, standard deviation of the mean, and the ratio of the standard deviation of the mean to the mean for each treatment. As many as 4,000 observations on as many as five input variables may be read into the machine. The analysis of variance is performed on the variable occupying the first position.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ EXECUTE Card
- (3) Binary Program Deck

- (4) Δ Card
- (5) Problem Card
- (6) Scale and Treatment Card(s)
- (7) Format Card(s)
- (8) Transgeneration Card(s) (if any)
- (9) Data

Items 5, 6, 7, 8 and 9 may be repeated as often as desired.

Master Control System Card

Starting in Column 1 the following characters appear:

Δ COOP, REQUEST NUMBER, NAME,, G.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	PR	Problem
3		Number of Input Variables, IM, $1 \leq IM \leq 5$
4-7		Total Number of Observations, M, $M \leq 4,000$
8,9		Number of Treatments, L, $2 \leq L < 100$
10		Number of Format Cards, NF, $1 \leq NF \leq 5$
11,12		Number of Transgeneration Cards, NTR, $NTR \leq 99$
13,72		Alphanumeric Problem Identification

Scale and Treatment Card(s)

<u>Columns</u>	<u>Explanation</u>
1-4	Scale Factor for Input Variable 1.
⋮	
16-20	" " " " " 5.
21-24	Number of Observations in Treatment 1.
25-28	" " " " " 2.
⋮	
69-72	" " " " " 13.

L + 5 four-column fields are necessary. Additional cards in the same format (19I4) are used as required.

The number punched in the Scale and Treatment Card(s) are signed integers (+ need not be punched). In the Scale part of the card the integers indicate the power of ten by which the input variable is to be multiplied. If the field is blank no scaling will be performed. The scale factors on the Scale Card must follow the same order in which the input variables are punched on the data cards, i.e. the first scale factor is for the first variable and so forth.

Format Card(s)

Columns 1-72 of up to five cards may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats. A format for the data must be specified.

Transgeneration Cards (if any)

Transgeneration Cards are explained in Appendix B. It is necessary that the number of Transgeneration Cards used agrees with the number punched in columns 11 and 12 of the Problem Card.

Data Cards

The Data Cards are punched according to the format on the Format Card(s).

EXAMPLE

In the example for ANOVA the input variables needed to be transformed to analyze the model

$$e^{(1.35395 + .061085X - .0056063X^2 + .0000025655X^3)} .$$

The transgeneration cards as shown performed the transformation. There are two treatments with 28 observations in each treatment.

^C00F,10168,NORMA HULL.
^EXECUTE.

BINARY PROGRAM DECK

^
PRI 56 21110P24G S 40.47

28 28

(BX,13F4.2/BX,15F4.2/BX,14F4.2/BX,14F4.2)

TR 2= 1TIMES 1

TR 3= 2TIMES 1

TR 3= 3TIMES .0025655

TR 3= 3TIMES .001

TR 2= 2TIMES .0056663

TR 2= 2TIMES .1

TR 1= 1TIMES 0.061085

TR 1= 1PLUS 1.35395

TR 1= 1PLUS 2

TR 1= 1PLUS 3

TR 1= EXP 1

BP24G9406115599361625994603960926044600061106100619560956111

BP24G940609960596099601960806109612061606119612961196116612962516150

BP24G94746404510459646304595459945504550457044004412461445294650

BP24G94745204539442045324687449544914530448945404509453145994460

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

WEIGHTED ONE-WAY ANALYSIS OF VARIANCE

OP24G S 40,47

TRANSGENERATIONS

2 = 1 TIMES 1
 3 = 2 TIMES 1
 3 = 3 TIMES .0025655
 3 = 3 TIMES .001
 2 = 2 TIMES .0056063
 2 = 2 TIMES .1
 1 = 1 TIMES 0.061085
 1 = 1 PLUS 1.35395
 1 = 1 PLUS 2
 1 = 1 PLUS 3
 1 = EXP 1

WEIGHTED ANALYSIS OF VARIANCE

SOURCE	D.F.	SUMS OF SQUARES	MEAN SQUARES	F RATIO
MEAN	1	5.1388843541e+04		
AMONG TREATMENTS	1	1.6656601104e+03	1.6656601104e+03	8.0291207456e+03
WITHIN TREATMENTS	54	1.1202427864e+01	2.0745236785e-01	
TOTAL	55	1.6768625383e+03		

ESTIMATED WITHIN SIGMA 4.5546939287e-01

K 2.8000000000e+01

ESTIMATED AMONG VARIANCE 5.9480452075e+01

ESTIMATED AMONG SIGMA 7.7123571024e+00

GRAND MEAN 3.0292869134e+01

CHI SQUARE(1) 2.0791584626e-02

CORRECTED CHI SQUARE(1) 2.0431245204e-02

TREATMENTS	N	WEIGHTED MEAN	VARIANCE OF MEAN	STD DEV OF MEAN	STD DEV/MEAN	STD DEV
1	28	3.5746668776e+01	7.2034641547e-03	8.4873224015e-02	1.2563601107e-02	4.4910688741e-01
2	28	2.4839069492e+01	7.6145614896e-03	8.7261454775e-02	1.8589432945e-02	4.6174421675e-01

ANOVAGR

Norma C. Hull

(FORTRAN 63)

DESCRIPTION

This program is a modification of program ANOVA written by F. L. Miller, Jr. ANOVAGR performs a weighted analysis of variance for a one-way classification, i.e. a "between-within" analysis. The program handles unequal treatment sample sizes. Transgeneration of the data is allowed. Several problems may be run, one behind another.

The input includes the number of input variables, the number of observations, the number of treatments, and the observations from the experiment.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ Execute Card
- (3) Binary Program Deck
- (4) Δ Card
- (5) Problem Card
- (6) Treatment Card(s)
- (7) Format Card(s)
- (8) Transgeneration Card(s) (if any)
- (9) Data
- (10) Graph Set Up Card

- (11) Independent Variable Number Card
- (12) Graph Title Card(s)
- (13) End of Problem Card

Master Control System Card

Starting in column 1 the following characters appear:

Δ COOP, Request Number, Name, ϕ /HD2, G.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2		Problem
3,4		Number of Input Variables, IM, $1 \leq IM \leq 5$
5-7		Total Number of Observations, M, $M \leq 500$
8,9		Number of Treatments, L, $2 \leq L < 100$
10		Number of Format Cards, NF, $1 \leq NF \leq 5$
11,12		Number of Transgeneration Cards, NTR, $NTR \leq 99$
13-72		Alphanumeric Problem Identification

Treatment Card(s)

<u>Columns</u>	<u>Explanation</u>
1-4	Number of Observations in Treatment 1.
5-8	" " " " " 2.
⋮	
69-72	" " " " " 18.

Additional cards in the same format (18I4) are used as required.

Format Card(s)

Columns 1-72 of up to five cards may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats. A format for the data must be specified.

Transgeneration Cards (if any)

Transgeneration Cards are explained in Appendix B. It is necessary that the number of Transgeneration Cards used agree with the number punched in columns 11 and 12 of the Problem Card.

Data Cards

The Data Cards are punched according to the format on the Format Card(s). See Appendix C for details concerning formats.

Graph Cards

For each choice of the dependent variable (variance, variance of mean, or standard error of mean), the independent variable may be the mean or any of the transgenerated or input variables (if the input variable is not destroyed by transgeneration). If the mean is used as the only independent variable for a particular choice of a dependent variable, only one card, the Graph Set Up Card, need be used for the graph options. If, however, any of the input or transgenerated variables are used, the Graph Set Up Card, the Independent Variable Number Card, and as many Graph Title Cards must be used as there are input or trans-generated variables used as independent variables.

Graph Set Up Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1	1	If dependent variable is variance of mean
	2	" " " " variance
	3	" " " " Standard Error of mean
2	0	If independent variable is <u>not</u> mean
	1	" " " " mean
3-5		Number of graphs of independent variables <u>other than mean</u>
6-69		Title if mean is used as an independent variable

Independent Variable Number CardColumns

1,2	Number of the variable to be treated as independent variable 1.									
3,4	"	"	"	"	"	"	"	"	"	2.
⋮										
59-60	"	"	"	"	"	"	"	"	"	30.

There must be as many fields (I2) used as are indicated in columns 3-5 on the Graph Set Up Card. The number of independent variables other than the mean is limited to 30. Therefore, there is a maximum of 31 graphs for each choice of a dependent variable.

Title Cards for Graphs

Columns 1-64 may be used as alphanumeric information in identifying graphs of independent variables other than the mean. The title cards are in the same sequence in which the independent variables they identify

are ordered on the Independent Variable Number Card. For example, the first Title Card would be for the first independent variable punched in columns 1 and 2 on the Independent Variable Number Card while the fifth Title Card would identify the graph produced by the independent variable as indicated in columns 9 and 10 on the Independent Variable Number Card. The total number of Title Cards must agree with the number punched in columns 3-5 on the Graph Set Up Card.

End of Problem Card

If one problem is run behind another, an End of Problem Card must be inserted at the end of each problem. In columns 6-11, the characters END PR must be punched. It is essential that this card be used.

EXAMPLE

In the example for ANOVAGR the model

$$Y = \text{Radioactive SR (Radioactive CA X.0000012)}$$

is needed. To obtain the radioactive elements SR and CA the data must be transformed by combining the input variables according to the directions of the experimenter.

The 1604A job card must show that tape drive "2" is used for plotting.

PR14 821814SNAIL DATA DATA Y=RADIO SR/(RADIO CA * .0000012) ARR545

5 4 5 4 4 5 5 4 5 4 5 5 5 5 3 4
(F6.4,F7.4,F5.0,F2.0,F3.1,F2F4.3,F5.4,F6.0,F2.0,F5.5,2F4.3,F5.4)

TR 3= 3MINUS

4

TR 3= 3OVER

5

TR 3= 3TIMES

6

TR 4= 3TIMES

7

TR 3= 4TIMES

8

TR 9= 9MINUS

10

TR 9= 9OVER

5

TR 5= 4TIMES

11

TR 9= 9MINUS

5

TR 9= 9TIMES

12

TR 9= 9TIMES

13

TR 9= 9TIMES

14

TR 9= 9TIMES .0000012

TR 1= 3OVER

9

95977	761484	252835	82321	13242	8768	27941	4803	75321	901239	16067	111
95977	761484	233135	90321	13242	8768	261674	803	75321	901239	16067	121
95977	761484	235735	14321	13242	8768	224904	803	75321	901239	16067	131
95977	761484	133035	88321	13242	8768	125754	803	75321	901239	16067	141
95977	761484	165235	146321	13242	8768	183964	803	75321	901239	16067	151
27570	735366	200335	166321	13242	8768	201734	803	75321	901239	16067	211
27570	735366	361635	204321	13242	8768	390844	803	75321	901239	16067	221
27570	735366	141735	18321	13242	8768	127864	803	75321	901239	16067	231
27570	735366	227935	19321	13242	8768	261864	803	75321	901239	16067	251
29918	321735	137435	13321	13242	8768	154874	803	75321	901239	16067	311
29918	321735	388835	18321	13242	8768	467134	803	75321	901239	16067	321
29918	321735	582435	18321	13242	8768	712564	803	75321	901239	16067	331
29918	321735	441835	140321	13242	8768	552084	803	75321	901239	16067	341
29918	321735	505335	18321	13242	8768	621234	803	75321	901239	16067	351
55003	91530277	112635	18321	13242	8768	126064	803	75321	901239	16067	411
55003	91530277	212035	150321	13242	8768	256204	803	75321	901239	16067	421
55003	91530277	80435	130321	13242	8768	881148	803	75321	901239	16067	431
55003	91530277	79235	164321	13242	8768	788848	803	75321	901239	16067	451
167292	1503	783635	120321	13242	8768	679174	803	75321	901239	16067	511
167292	1503	532735	146321	13242	8768	693194	803	75321	901239	16067	521
167292	1503	615735	124321	13242	8768	772804	803	75321	901239	16067	531
167292	1503	895635	128321	13242	8768	1043864	803	75321	901239	16067	541
3841232	705832	99735	106321	13242	8768	100264	803	75321	901239	16067	611
3841232	705832	133335	156321	13242	8768	130354	803	75321	901239	16067	621
3841232	705832	145735	156321	13242	8768	145954	803	75321	901239	16067	631
3841232	705832	162635	140321	13242	8768	161294	803	75321	901239	16067	641
3841232	705832	115135	174321	13242	8768	121354	803	75321	901239	16067	651
455437	945852	248235	18321	13242	8768	280574	803	75321	901239	16067	711
455437	945852	179835	162321	13242	8768	192524	803	75321	901239	16067	721
455437	945852	415935	174321	13242	8768	436524	803	75321	901239	16067	731
455437	945852	288535	152321	13242	8768	359844	803	75321	901239	16067	741
455437	945852	303435	194321	13242	8768	357564	803	75321	901239	16067	751
27570	735366	200335	164321	13242	8768	230324	803	75321	901239	16067	811
27570	735366	111535	122321	13242	8768	110484	803	75321	901239	16067	821
27570	735366	262935	122321	13242	8768	315524	803	75321	901239	16067	831
27570	735366	113035	160321	13242	8768	101724	803	75321	901239	16067	841
27570	735366	153235	156321	13242	8768	164894	803	75321	901239	16067	851
119451	729551	467535	144321	13242	8768	590714	803	75321	901239	16067	911
119451	729551	370035	164321	13242	8768	478844	803	75321	901239	16067	921
119451	729551	232335	152321	13242	8768	270044	803	75321	901239	16067	931
119451	729551	305535	120321	13242	8768	330954	803	75321	901239	16067	941
27570	735366	245734	176321	13402	8768	236464	803	75321	901239	16067	1011
27570	735366	281534	254321	13402	8768	305534	803	75321	901239	16067	1021


```

2757071353668 3816342083215134028768 4391748037532190123916067 1031
2757071353668 3885341423215134028768 4768648037532190123916067 1041
2757071353668 1356341403215134028768 1251048037532190123916067 1051
2757071353668 1938341823215134028768 2392748037532190123916067 1111
2757071353668 905341203215134028768 858348037532190123916067 1121
2757071353668 924341743215134028768 1076648037532190123916067 1131
2757071353668 1233341623215134028768 1306048037532190123916067 1141
2757071353668 1996341243215134028768 1953548037532190123916067 1211
2757071353668 1675341663215134028768 1772948037532190123916067 1221
2757071353668 944341463215134028768 964648037532190123916067 1231
2757071353668 1042342263215134028768 1213748037532190123916067 1241
2757071353668 1796341843215134028768 1492648037532190123916067 1251
4319631977784 1234341403215134028768 1271148037532190123916067 1311
4319631977784 1263341423215134028768 1252848037532190123916067 1321
4319631977784 2113341533215134028768 2214448037532190123916067 1331
4319631977784 2147341563215134028768 2278548037532190123916067 1341
4319631977784 1785342043215134028768 1949848037532190123916067 1351
2757071353668 1719341643215134028768 1938148037532190123916067 1411
2757071353668 1619341703215134028768 1743548037532190123916067 1421
2757071353668 1527341063215134028768 2260648037532190123916067 1431
2757071353668 1490341963215134028768 1456248037532190123916067 1441
2757071353668 2702342043215134028768 3098248037532190123916067 1451
2522332385600 1449341203215134028768 1448748037532190123916067 1511
2522332385600 1064341643215134028768 913948037532190123916067 1521
2522332385600 1740341583215134028768 1837748037532190123916067 1531
2522332385600 1568341543215134028768 1671248037532190123916067 1541
2522332385600 1226341223215134028768 1233648037532190123916067 1551
441624178113 7952341683215134028768 10341448037532190123916067 1611
441624178113 7266341063215134028768 8920348037532190123916067 1621
441624178113 6922341743215134028768 8985648037532190123916067 1631
441624178113 9778341463215134028768 12150848037532190123916067 1641
441624178113 5406341993215134028768 2000448037532190123916067 1651
1375177058 1601341283215134028768 1231248037532190123916067 1711
1375177058 1883341163215134028768 1743048037532190123916067 1731
1375177058 1120341323215134028768 1077048037532190123916067 1741
1097912529222 1180341473215134028768 1627848037532190123916067 1811
1097912529222 1681341513215134028768 1714348037532190123916067 1821
1097912529222 1272341463215134028768 1338448037532190123916067 1831
1097912529222 1757341623215134028768 1872548037532190123916067 1951
21 2 IS VARIANCE, X IS MEAN

```

```

Y2 IS VARIANCE, X IS CA INPUT
Y IS VARIANCE, X IS SR INPUT
31 2 Y IS STANDARD DEVIATION OF MEAN, X IS MEAN
Y2 IS STANDARD DEVIATION OF MEAN, X IS CA INPUT
Y IS STANDARD DEVIATION OF MEAN, X IS SR INPUT

```

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

WEIGHTED ONE-WAY ANALYSIS OF VARIANCE WITH GRAPHING OPTIONS

SNAIL DATA DATA Y=RADIO SR/(RADIO CA * .0000012) ARROWS

TRANSGENERATIONS

3 = 3 MINUS 4
3 = 3 OVER 5
3 = 3 TIMES 6
4 = 3 TIMES 7
3 = 4 TIMES 8
9 = 9 MINUS 10
9 = 9 OVER 5
9 = 4 TIMES 11
9 = 9 MINUS 5
9 = 9 TIMES 12
9 = 9 TIMES 13
9 = 9 TIMES 14
9 = 9 TIMES .0000012
1 = 3 OVER 9

WEIGHTED ANALYSIS OF VARIANCE

SOURCE	D.F.	SUMS OF SQUARES	MEAN SQUARES	F RATIO
MEAN	1	3.8984939543E+12		
AMONG TREATMENTS	17	2.5028270720E+10	1.4722512188E+09	2.6993064548E+00
WITHIN TREATMENTS	64	3.4906772992E+10	5.4541832799E+08	
TOTAL	81	5.9935043711E+10		

ESTIMATED WITHIN SIGMA 2.3354192942E+04

K 4.3509325681E+00

ESTIMATED AMONG VARIANCE 2.0365779473E+08

ESTIMATED AMONG SIGMA 1.4270872248E+04

GRAND MEAN 2.1776284802E+05

CHI SQUARE(17) 3.6913889110E+01

CORRECTED CHI SQUARE(17) 3.4218225397E+01

TREATMENTS	N	WEIGHTED MEAN	VARIANCE	VARIANCE OF MEAN	STD DEV OF MEAN	STD DEV/MEAN
1	5	2.2420690188E+05	3.9165226800E+05	7.8330453599E+07	8.8504493444E+03	3.9474473221E+02
2	4	2.2786904809E+05	5.4401944665E+06	1.3600486156E+08	1.1662112230E+04	5.1179009731E+02
3	5	1.9415590854E+05	4.4549347999E+07	8.9098695996E+06	2.9849404683E+03	1.5373935775E+02

4	4	2.0910157643**05	2.2421751466**08	5.6054378666**07	7.4869472194**03	3.5805312170**02
5	4	2.1121643073**05	1.8200961920**09	4.5502404900**08	2.1331292694**04	1.0099258196**01
6	5	2.3051578858**05	5.1035164000**07	1.0207032300**07	3.1948447223**03	1.3859548371**02
7	5	2.0724434898**05	2.1393429600**08	4.2786859198**07	6.5411665014**03	3.1562577711**02
8	5	2.2093695559**05	6.2055893199**08	1.2411786400**08	1.1140546952**04	5.0424099138**02
9	4	1.9635005081**05	2.5721908533**08	6.4304771333**07	8.0190255849**03	4.0840455866**02
10	5	2.2468194882**05	6.8820237300**08	1.3764047450**08	1.1732027727**04	5.2216151517**02
11	4	2.1453985993**05	5.9816484399**08	1.4954121100**08	1.2228704387**04	5.6999684772**02
12	5	2.3601075315**05	1.0126592620**09	2.0253785239**08	1.4231368606**04	6.0299659284**02
13	5	2.2626639512**05	5.6037637000**07	1.1207527400**07	3.3477645377**03	1.4799677174**02
14	5	2.0717600727**05	9.2568399899**08	1.8513679990**08	1.3606498440**04	6.5676033720**02
15	5	2.3760270560**05	4.3862622599**08	8.7722245198**07	9.3661755905**03	3.9419482060**02
16	5	1.8788020879**05	2.8233175999**07	5.6456351999**06	2.3762649683**03	1.2647766274**02
17	3	2.6981025504**05	1.1952659340**09	3.9842497900**08	1.9960510464**04	7.3979806515**02
18	4	2.0685626172**05	1.4334664240**09	3.5836660599**08	1.8930573314**04	9.1515592306**02

BMD02VH
Norma C. Hull
(Fortran '63)

DESCRIPTION

This program is a modification of the program "BMD02V" written at the University of California at Los Angeles. BMD02VH will produce an analysis of variance for a factorial experiment. For each problem the number of variables must be less than or equal to eight; the number of replicates must be less than or equal to 999; the number of categories or levels of any one variable must be less than or equal to 999 and the product of the categories or levels must be less than or equal to 18,000; and the number of variable format cards must be less than or equal to five. The program can perform transgenerations on input data, either according to the codes specified on one Special Transgeneration Card or according to the codes on the Standard Transgeneration Cards. The Special Transgeneration List and the Standard Transgeneration List are peculiar to this program. (See Page

The input includes the number of observations per treatment, the number of factors, the variable format, the number of factors for which an orthogonal polynomial breakdown is desired, and the data from the experiment.

The output includes an analysis-of-variance table, the grand mean, a breakdown of the sums of squares into orthogonal polynomial components for the maximum number of main effects (8) and all of their first order interactions, and cell and marginal means.

CARD PREPARATION

Order of Cards

- (1) Master Control System Card
- (2) Δ Execute Card
- (3) Binary Program Deck
- (4) Δ Card
- (5) Problem Card
- (6) Title Card
- (7) Variable Name Card
- (8) Special Transgeneration Card (if used)
- (9) Standard Transgeneration Card(s) (if used)
- (10) Variable Format Card(s)
- (11) Data Input Cards
- (12) Finish Card.

Items 5 through 11 may be repeated as often as desired.

Master Control System Card

Starting in column 1 the following characters appear:

Δ COOP, Request Number, Name,, G.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1-6	PROBLM	(mandatory)
7-8		Problem number (alphanumeric)
9-11		Number of observations per cell ($R \leq 999$)

- 12 Number of factors
($W \leq 8$)
- 13 1 If Special Transgeneration Card follows;
blank otherwise.
- 14 2 If cell and marginal means are to be printed;
1 If only marginal means are to be printed;
If no means are desired, leave the column blank.
- 15,16 Number of input data fields per card. See the
Data Input Card Section below.
- 17 Number of variables for which an orthogonal
polynomial breakdown is desired (≤ 8). Only
variable with nine or fewer levels may be
used in this breakdown. The variables are
specified in ascending order in the following
eight columns. If no breakdown is desired,
leave columns 17-21 blank.
- 18 1st variable for orthogonal polynomial breakdown
- 19 2nd " " " " "
- 20 3rd " " " " "
- 21 4th " " " " "
- 22 5th " " " " "
- 23 6th " " " " "
- 24 7th " " " " "
- 25 8th " " " " "
- 26-28 Number of levels of the 1st factor
- 29-31 " " " " " 2nd "
- 32-34 " " " " " 3rd "
- 35-37 " " " " " 4th "
- 38-40 " " " " " 5th "
- 41-43 " " " " " 6th "
- 44-46 " " " " " 7th "
- 47-49 " " " " " 8th "

50-51	Number of Standard Transgeneration Cards
52-53	Number of Variables per observation
54-68	Blank
69-70	If data input is from a tape other than the standard input tape, punch that tape unit number; otherwise leave blank.
71,72	Number of Variable Format Cards ($1 \leq k \leq 5$)

Title Card

Columns 1-80 may be used for any alphanumeric problem identification.

This card must be included.

Variable Name Card

<u>Columns</u>	<u>Explanation</u>
1-8	Any alphanumeric identification for variable 1
9-16	" " " " " 2
17-24	" " " " " 3
25-32	" " " " " 4
33-40	" " " " " 5
41-48	" " " " " 6
49-56	" " " " " 7
57-64	" " " " " 8

This card must be included.

Special Transgeneration Card

This card is used only in those programs which require transformation of the form $f(Y) \rightarrow Y$. One "Special" Transgeneration Card will specify successive transformations.

<u>Columns</u>	<u>Explanation</u>
1-6	SPECTG (Mandatory)
7	Number of transformations (≤ 8 for each card)
8,9	Code* for the 1st transgeneration
10-15	Constant** for the 1st transgeneration (if none, leave blank)
16,17	Code* for the 2nd transgeneration
18-23	Constant** for the 2nd transgeneration
..	
64,65	Code* for the 8th transgeneration
66,71	Constant** for the 8th transgeneration

As an example, if the user desires the transformation

$$(\sqrt{Y^2} - 3 + 100)^2 \rightarrow Y$$

he could accomplish this with the Special Transgeneration Card

SPECTG5102.000008-3.00001000000008100.00102.0000

*Codes 01, 02...,10 of the Transgeneration List may be used.

**Keypunch decimal point. The decimal point is not required for right-justified integers.

Standard Transgeneration Card(s)

If a non-zero number is specified in Col. 50 and 51 of the Problem Card, the same number of Transgeneration Card(s) must be prepared. Different types of transformation can be performed successively on the same variable, if desired. The format of the Standard Transgeneration Card is as follows.

<u>Columns</u>	<u>Explanation</u>
1,2	TR (for TRansgeneration Card)
3,4	Factor number to be assigned on transformed or generated factor
5,6	Transformation code*
7,8	X_i factor number
9-14	X_j factor number or constant (keypunch with decimal point). If the constant is a negative value, keypunch a minus sign to the left of the constant.

*Codes 01, 02, ..., 23 of the Transgeneration List may be used.

Variable Format Card(s)

Columns 1-72 of up to five cards may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats.

Data Input Cards

The form of the data input is illustrated in the following example, which consists of a three-factor design where two factors have three levels, the other has two levels, and these are two

observations at each treatment combination. The data may be represented in the form:

$$x_{rijk} \quad \begin{array}{l} r = 1, 2 \\ i = 1, 2, 3 \\ j = 1, 2, 3 \\ k = 1, 2 \end{array}$$

The program considers

i as the index for the first factor
 j " " " " " second "
 k " " " " " third "

The data are lexicographically ordered by their subscripts. That is, the subscript on the right, namely factor 3, is incremented first, then factor 2, then factor 1, and finally the observations at the treatment combination.

x_{r111}

x_{r112}

x_{r121}

x_{r122}

x_{r131}

x_{r132}

x_{r211}

$\vdots$

x_{r332}

When the subscript r is incremented the observations must be started on a new card. Any unused fields on the previous card may be left blank. This order must be maintained since it is the only means of identification

of the data for the program. Each Data Input Card must have the same format, and this format must be specified on the F-type Variable Format Card. The number of data fields per card must be specified in Columns 15, 16 of the Problem Card.

Finish Card

FINISH in Col. 1-6 is placed after the last problem to be run for a job.

COMPUTATIONAL PROCEDURE

Step 1. The calculations performed by this program will be demonstrated, using as an example a three-way analysis of variance with R observations per cell.

Let x_{rijk} $r = 1, 2, \dots, R$
 $i = 1, 2, \dots, I$
 $j = 1, 2, \dots, J$
 $k = 1, 2, \dots, K$

represent the data after transgeneration.

Step 2. The first item computed is the grand mean $\bar{x} \dots$

The TOTAL and the WITHIN sums of squares are computed:

$$\text{TOTAL} = \sum_r \sum_i \sum_j \sum_k (x_{rijk} - \bar{x} \dots)^2$$

$$\text{WITHIN} = \sum_r \sum_i \sum_j \sum_k (x_{rijk} - \bar{x}_{.ijk})^2$$

and cell means, $\bar{x}_{.ijk}$ are computed.

Step 3. The grand mean is subtracted from these to obtain

$$(\bar{x}_{.ijk} - \bar{x} \dots) \begin{matrix} i = 1, 2, \dots, I \\ j = 1, 2, \dots, J \\ k = 1, 2, \dots, K \end{matrix}$$

Step 4. Sums of squares are then computed using formulas of the type:

$$A = \frac{R}{IJK} \left\{ I \sum_i \left(\sum_j \sum_k (\bar{x}_{.ijk} - \bar{x} \dots) \right)^2 \right\}$$

$$B = \frac{R}{IJK} \left\{ J \sum_j \left(\sum_i \sum_k (\bar{x}_{.ijk} - \bar{x} \dots) \right)^2 \right\}$$

$$C = \frac{R}{IJK} \left\{ K \sum_k \left(\sum_i \sum_j (\bar{x}_{.ijk} - \bar{x} \dots) \right)^2 \right\}$$

$$AB = \frac{R}{IJK} \left\{ I \sum_i \left(\sum_j \sum_k (\bar{x}_{.ijk} - \bar{x} \dots) \right)^2 \right\} - (A+B)$$

$$AC = \frac{R}{IJK} \left\{ I \sum_i \left(\sum_k \sum_j (\bar{x}_{.ijk} - \bar{x} \dots) \right)^2 \right\} - (A+C)$$

Step 5. $BC = \frac{R}{IJK} \left\{ J \sum_j \left(\sum_k \sum_i (\bar{x}_{.ijk} - \bar{x} \dots) \right)^2 \right\} - (B+C)$

$$ABC = \text{TOTAL} - \text{WITHIN} - (A + B + C + AB + AC + BC)$$

If requested on the Problem Card, a breakdown of certain main effect and first order interaction sums of squares into orthogonal polynomial components is made. The nature of these computations is explained in the following example.

The first and third factors of the previous example are designated for this analysis. (This means I and K are less than or equal to 9.)

$$\begin{aligned} \text{Let } x_{i++} &= \sum_j \sum_k \bar{x}_{.ijk} \\ x_{++k} &= \sum_i \sum_j \bar{x}_{.ijk} \\ x_{i+k} &= \sum_j \bar{x}_{.ijk} \end{aligned}$$

The following matrices are printed:

$$X_i = \begin{vmatrix} x_{1++} \\ x_{2++} \\ \vdots \\ x_{I++} \end{vmatrix}, X_k = \begin{vmatrix} x_{++1} \\ x_{++2} \\ \vdots \\ x_{++K} \end{vmatrix}, X_{ij} = \begin{vmatrix} x_{1+1} & x_{1+2} & \cdots & x_{1+K} \\ x_{2+1} & x_{2+2} & \cdots & x_{2+K} \\ \vdots & \vdots & & \vdots \\ x_{I+1} & x_{I+2} & \cdots & x_{I+K} \end{vmatrix}$$

Let

$$P_l = \begin{vmatrix} P_{11} & P_{12} & \cdots & P_{1l} \\ P_{21} & P_{22} & \cdots & P_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ P_{l-1,1} & P_{l-1,2} & \cdots & P_{l-1,l} \end{vmatrix}$$

where p_{ij} is the j th value of the orthogonal polynomial of degree i defined in l equally spaced points. The following matrix multiplications are performed:

$$P_I X_i, P_K X_K, P_I X_{ik} P_K$$

When the elements of these matrices are squared and divided respectively by the constants $J \cdot K/R$, $I \cdot J/R$, and J/R , the resulting matrices are printed. They represent respectively the orthogonal polynomial components of the sums of squares for the i main effect, the k main effect, and the $i \times k$ interaction. In the case when these matrices have more than three rows or columns (i.e., when I or $K > 4$), all the rows beyond the first three are replaced by their sum and then all the columns of the resulting matrices beyond the first three are replaced by their sum. These final matrices are printed under the row and column headings, LINEAR, QUADRATIC, CUBIC REMAINDER.

Step 6. For every pair of variables treated in Step 5, a table of main effect and interaction means is printed. For the example given in Step 5, this table would be

$$\begin{array}{|c|c|} \hline \bar{X}_{.i.k} - \bar{X}_{.i..} - \bar{X}_{...k} + \bar{X}_{....} & I \times K \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline \bar{X}_{.i..} - \bar{X}_{....} & I \times 1 \\ \hline \end{array}$$

$$\begin{array}{|c|c|} \hline \bar{X}_{...k} - \bar{X}_{....} & 1 \times K \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline \bar{X}_{....} & 1 \times 1 \\ \hline \end{array}$$

Step 7. For every pair of factors treated in Step 5, a table of means is printed. For the example in Step 5, this table would be

$$\begin{array}{|c|c|} \hline \bar{X}_{.i.k} & 1 \times K \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline \bar{X}_{.i..} & 1 \times 1 \\ \hline \end{array}$$

$$\begin{array}{|c|c|} \hline \bar{X}_{...k} & 1 \times K \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline \bar{X}_{....} & 1 \times 1 \\ \hline \end{array}$$

Step 8. If requested on the Problem Card, cell and marginal means are printed. For the example in Step 5, these include

$$\bar{x}_{.ijk} \quad \bar{x}_{.i..} \quad \bar{x}_{..jk} \quad \bar{x}_{...k}$$

for $i = 1, \dots, I$; $j = 1, \dots, J$; $k = 1, \dots, K$

TRANSGENERATION CARDS

The term transgeneration is used to include transformations of input variables and creation of new variables prior to the normal computations performed by the program.

Transgeneration List

Notation to be used in the following transgeneration list:

i, j, k are variable indices (need not be different)

c is a constant

$a_1, a_2, a_3, \dots$ are constants

n is the number of cases, or sample sizes

<u>Code</u>	<u>Transgeneration</u>	<u>Restriction</u>
01	$\sqrt{X_i} \rightarrow X_k$	$X_i \geq 0$
02	$\sqrt{X_i} + \sqrt{X_i} + 1 \rightarrow X_k$	$X_i \geq 0$
03	$\log_{10} X_i \rightarrow X_k$	$X_i > 0$
04	$e^{X_i} \rightarrow X_k$	-
05	$\sin^{-1} \sqrt{X_i} \rightarrow X_k$	$0 \leq X_i \leq 1$
06	$\sin^{-1} \sqrt{X_i/(n+1)} + \sin^{-1} \sqrt{(X_i+1)/(n+1)} \rightarrow X_k$	$0 \leq (X_i/n) \leq 1$

07	$1/X_i \rightarrow X_k$	$X_i \neq 0$
08	$X_i + c \rightarrow X_k$	-
09	$X_i c \rightarrow X_k$	-
10	$X_i^c \rightarrow X_k$	$X_i \geq 0$
11	$X_i + X_j \rightarrow X_k$	-
12	$X_i - X_j \rightarrow X_k$	-
13	$X_i X_j \rightarrow X_k$	-
14	$X_i/X_j \rightarrow X_k$	$X_j \neq 0$
15	$\log_e X_i \rightarrow X_k$	$X_i > 0$
16	$\sin X_i \rightarrow X_k$	-
17	$\cos X_i \rightarrow X_k$	-
18	$\tan^{-1} X_i \rightarrow X_k$	-
19	$X_i^{X_j} \rightarrow X_k$	$X_i > 0$
20	$c^{X_i} \rightarrow X_k$	$c > 0$
21	$ X_i \rightarrow X_k$	
22	$X_i \rightarrow \text{Truncation, integer } (X_k)$	
23	$X_i \text{ modulo } c \rightarrow X_k$	

Examples:

$$\log_{10} X_4 \rightarrow X_4$$

$\log_{10} X_4$ replaces X_4

$$X_5^c \rightarrow X_1$$

X_5^c replaces X_1

By successive transformations, more complicated relationships may be obtained. For example:

- (i) To replace X_5 by $\sqrt{X_1^2 + X_3^2}$ four transformations are required:

Variables as they are stored at each step

Transformation	X_1	X_2	X_3	X_4	X_5
$X_1^2 \rightarrow X_1$	X_1^2	X_2	X_3	X_4	X_5
$X_3^2 \rightarrow X_3$	X_1^2	X_2	X_3^2	X_4	X_5
$X_1 + X_3 \rightarrow X_5$	X_1^2	X_2	X_3^2	X_4	$X_1^2 + X_3^2$
$\sqrt{X_5} \rightarrow X_5$	X_1^2	X_2	X_3^2	X_4	$\sqrt{X_1^2 + X_3^2}$

In this example, it can be seen that the original values of X_5 are irrelevant. Actually the variable X_5 may be a dummy variable introduced by the program specifically to provide capacity for creating new variables by transgeneration. Dummy variables may be required for intermediate storage in order to effect some transformations.

- (ii) To replace X_1 by $\exp(-1/2X_1^2)$ three transformations are required:

Transformation	X_1	X_2	X_3
$X_1^2 \rightarrow X_1$	X_1^2	X_2	X_3
$-1/2X_1 \rightarrow X_1$	$-1/2X_1^2$	X_2	X_3
$\exp(X_1) \rightarrow X_1$	$\exp(-1/2X_1^2)$	X_2	X_3

EXAMPLE

In the example for BMD02VH there are five variables with levels of 3, 3, 3, 2, and 2. No transgenerations are desired. The tables of interactions are requested for all the variables. Note the example of the 1604A JOB CARD; tapes 2 and 3 must be assigned.

PROBLM 2 15 2 1512345 3 3 3 2 2
 CONGDON TEST DATA 90 DAY WITHOUT TRANSGENERATION
 AGE DAY DOSE SEX REPS
 (F13.0)

.32175055442 11111
 .46364760903 11112
 .32175055442 11121
 .32175055442 11122
 .32175055442 11211
 .36136712394 11212
 .32175055442 11221
 .32175055442 11222
 .46364760903 11311
 .32175055442 11312
 .32175055442 11321
 .52359877557 11322
 .46364760903 12111
 .46364760903 12112
 .88607712376 12121
 .32175055442 12122
 .32175055442 12211
 .46364760903 12212
 .46364760903 12221
 .32175055442 12222
 .36136712394 12311
 .46364760903 12312
 .32175055442 12321
 .36136712394 12322
 .46364760903 13111
 .32175055442 13112
 .88607712376 13121
 .32175055442 13122
 .1071487178 13211
 .46364760903 13212
 .68471920297 13221
 .29284277175 13222
 .32175055442 13311
 .32175055442 13312
 .32175055442 13321
 .32175055442 13322
 .32175055442 21111
 .2490457724 21112
 .46364760903 21121
 .46364760903 21122
 .46364760903 21211
 .32175055442 21212
 .88607712376 21221
 .46364760903 21222
 .32175055442 21311
 .32175055442 21312
 .32175055442 21321
 .32175055442 21322
 .46364760903 22111
 .46364760903 22112
 .32175055442 22121
 .46364760903 22122
 .32175055442 22211
 .46364760903 22212
 .32175055442 22221
 .32175055442 22222

.68471920297 22311
 .46364760903 22312
 .32175055442 22321
 .46364760903 22322
 .32175055442 23111
 .32175055442 23112
 .46364760903 23121
 .46364760903 23122
 .32175055442 23211
 .32175055442 23212
 .36136712394 23221
 .32175055442 23222
 .52359877557 23311
 .68471920297 23312
 .32175055442 23321
 .46364760903 23322
 .46364760903 31111
 .78539816341 31112
 .32175055442 31121
 .68471920297 31122
 .32175055442 31211
 .46364760903 31212
 .32175055442 31221
 .46364760903 31222
 .46364760903 31311
 .32175055442 31312
 .2490457724 31321
 .46364760903 31322
 .88607712376 32111
 .36136712394 32112
 .68471920297 32121
 .32175055442 32122
 .32175055442 32211
 .46364760903 32212
 .46364760903 32221
 .68471920297 32222
 .88607712376 32311
 .46364760903 32312
 .68471920297 32321
 .68471920297 32322
 .32175055442 33111
 .32175055442 33112
 .46364760903 33121
 .32175055442 33122
 .32175055442 33211
 .46364760903 33212
 .32175055442 33221
 .32175055442 33222
 .32175055442 33311
 .80607712376 33312
 .32175055442 33321
 .11071487178 33322
 FINISH

BMD02V - ANALYSIS OF VARIANCE FOR FACTORIAL DESIGN - VERSION OF DECEMBER 10, 1965
 STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

CONGDON TEST DATA 90 DAY WITHOUT TRANSGENERATION

PROBLEM NO. 2

NUMBER OF VARIABLES 5
 NUMBER OF REPLICATES 1

VARIABLE	VARIABLE NAME	NO. OF LEVELS
1	AGE	3
2	DAY	3
3	DOSE	3
4	SEX	2
5	REPS	2

GRAND MEAN .46335

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUMS OF SQUARES	MEAN SQUARES
1	2	.17721	.08861
2	2	.01350	.00675
3	2	.08356	.04178
4	1	.00004	.00004
5	1	.00389	.00389
12	4	.20796	.05199
13	4	.39204	.09551
14	2	.05002	.02501
15	2	.12344	.06172
23	4	.09034	.02258
24	2	.01026	.00513
25	2	.04585	.02293
34	2	.00474	.00237
35	2	.01813	.00906
45	1	.03083	.03083
123	8	.43437	.05430
124	4	.02586	.00649
125	4	.34628	.08657
134	4	.22309	.05577
135	4	.22570	.05642
145	2	.01533	.00767
234	4	.30539	.07635
235	4	.79120	.19780
245	2	.04735	.02367
345	2	.09528	.04264
1234	8	.13502	.01688
1235	8	.35112	.04389
1245	4	.19031	.04508
1345	4	.06141	.01535

2345	4	.03018	.00754
RESIDUAL	8	.20255	.02532
TOTAL	107	4.70234	

ORDERED VARIABLES... 1 2 3 4 5

1	LINEAR	.15096	
1	QUADRATIC	.02625	
1	TOTAL		.17721

2	LINEAR	.00241	
2	QUADRATIC	.01109	
2	TOTAL		.01350

3	LINEAR	.00066	
3	QUADRATIC	.08291	
3	TOTAL		.08356

4	LINEAR	.00004	
4	TOTAL		.00004

5	LINEAR	.00389	
5	TOTAL		.00389

VARIABLE 1 ORDERED	VARIABLE 2 ORDERED				TOTAL
	LINEAR	QUADRATIC	CUBIC	REMAINDER	
LINEAR	.10756	.02204	0	0	.12960
QUADRATIC	.04953	.02883	0	0	.07836
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.15709	.05087	0	0	.20796

TABLE OF INTERACTIONS FOR

VARIABLE 1 (ROW)		VARIABLE 2 (COLUMN)	
.06160	.00840	.07007	.03477
.05343	.03268	.02075	.02203
.00825	.04108	.04933	.05681
.00130	.01413	.01295	.46376

TABLE OF MEANS

.36552	.43450	.48572	.42858
.49335	.42295	.40759	.44130
.52703	.57557	.45788	.52016
.46197	.47767	.45040	.46335

VARIABLE 1 ORDERED	VARIABLE 3 ORDERED				TOTAL
	LINEAR	QUADRATIC	CUBIC	REMAINDER	
LINEAR	.19342	.16226	0	0	.35569
QUADRATIC	.02556	.00080	0	0	.02636
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.21898	.16306	0	0	.38204

TABLE OF INTERACTIONS FOR

VARIABLE 1 (ROW)	VARIABLE 3 (COLUMN)			
.01795	.06442	-.08237	-.03477	
.02393	.00544	-.02937	-.02205	
-.04187	-.06986	.11173	.05681	
.01657	-.03918	.02261	.46335	

TABLE OF MEANS

.46310	.45381	.36882	.42858
.48180	.40755	.43454	.44130
.49486	.41112	.65450	.52016
.47992	.42416	.48595	.46335

VARIABLE 1 ORDERED	VARIABLE 4 ORDERED				TOTAL
	LINEAR	QUADRATIC	CUBIC	REMAINDER	
LINEAR	.02004	0	0	0	.02004
QUADRATIC	.02998	0	0	0	.02998
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.05002	0	0	0	.05002

TABLE OF INTERACTIONS FOR

VARIABLE 1 (ROW)	VARIABLE 4 (COLUMN)			
.00490	-.00490	-.03477		
.02356	-.02356	-.02205		
-.02846	.02846	.05681		
-.00063	.00063	.46335		

TABLE OF MEANS

.43285	.42431	.42858
.46423	.41837	.44130
.49106	.54926	.52016
.46271	.46398	.46335

VARIABLE 1 ORDERED	VARIABLE 5 ORDERED				TOTAL
	LINEAR	QUADRATIC	CUBIC	REMAINDER	

VARIABLE 1 ORDERED

LINEAR	.07777	0	0	0	.07777
QUADRATIC	.04567	0	0	0	.04567
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.12344	0	0	0	.12344

TABLE OF INTERACTIONS FOR

VARIABLE 1 (ROW)	VARIABLE 5 (COLUMN)	
.04741	-.04741	-.03477
-.02908	.02908	-.02205
-.01933	.01933	.05681
.00600	-.00600	.46335

TABLE OF MEANS

.48199	.37518	.42850
.41821	.46438	.44130
.50783	.53249	.52016
.46934	.45735	.46335

VARIABLE 3 ORDERED

VARIABLE 2 ORDERED	LINEAR	QUADRATIC	CUBIC	REMAINDER	TOTAL
LINEAR	.05959	.01052	0	-0	.07010
QUADRATIC	.00001	.02022	0	0	.02024
CUBIC	0	0	0	0	0
REMAINDER	-0	0	0	0	-0
TOTAL	.05960	.03074	0	-0	.09034

TABLE OF INTERACTIONS FOR

VARIABLE 2 (ROW)	VARIABLE 3 (COLUMN)	
.03666	-.00341	-.03326
.01423	-.02737	.01314
-.05089	.03078	.02012
.01657	-.03918	.02261
		.46335

TABLE OF MEANS

.51521	.41938	.45132	.46197
.50848	.41112	.51342	.47767
.41608	.44199	.49312	.45040
.47992	.42416	.48595	.46335

VARIABLE 4 ORDERED

VARIABLE 2 ORDERED	LINEAR	QUADRATIC	CUBIC	REMAINDER	TOTAL
LINEAR	.00395	0	0	0	.00395

QUADRATIC	.00631	0	0	0	.00631
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.01026	0	0	0	.01026

TABLE OF INTERACTIONS FOR

VARIABLE 2 (ROW)			VARIABLE 4 (COLUMN)		
-.01281	.01281	-.00138			
.01081	-.01081	.01433			
.00200	-.00200	.01295			
-.00063	.00063	.46335			

TABLE OF MEANS

.48853	.47541	.46197
.48785	.46750	.47767
.45176	.44903	.45040
.46271	.46398	.46335

VARIABLE 5 ORDERED

VARIABLE 2 ORDERED	LINEAR	QUADRATIC	CUBIC	REMAINDER	TOTAL
LINEAR	.00832	0	0	0	.00832
QUADRATIC	.03753	0	0	0	.03753
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.04585	0	0	0	.04585

TABLE OF INTERACTIONS FOR

VARIABLE 2 (ROW)			VARIABLE 5 (COLUMN)		
-.02393	.02393	-.00138			
.02636	-.02636	.01433			
-.00243	.00243	.01295			
.00600	-.00600	.46335			

TABLE OF MEANS

.44403	.47990	.46197
.51003	.44531	.47767
.45396	.44683	.45040
.46934	.45735	.46335

VARIABLE 4 ORDERED

VARIABLE 3 ORDERED	LINEAR	QUADRATIC	CUBIC	REMAINDER	TOTAL
LINEAR	.00471	0	0	0	.00471
QUADRATIC	.00003	0	0	0	.00003
CUBIC	0	0	0	0	0

REMAINDER	0	0	0	0	0
TOTAL	.00474	0	0	0	.00474

TABLE OF INTERACTIONS FOR

VARIABLE 3 (ROW)	VARIABLE 4 (COLUMN)	
.00847	-.00847	.01657
-.00076	.00076	-.03918
-.00771	.00771	.02261
-.00063	.00063	.46335

TABLE OF MEANS

.48776	.47208	.47992
.42277	.42556	.42416
.47761	.49429	.48595
.46271	.46398	.46335

VARIABLE 5 ORDERED					
VARIABLE 3 ORDERED	LINEAR	QUADRATIC	CUBIC	REMAINDER	TOTAL
LINEAR	.00957	0	0	0	.00957
QUADRATIC	.00856	0	0	0	.00856
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.01813	0	0	0	.01813

TABLE OF INTERACTIONS FOR

VARIABLE 3 (ROW)	VARIABLE 5 (COLUMN)	
.00523	-.00523	.01657
.01259	-.01259	-.03918
-.01782	.01782	.02261
.00600	-.00600	.46335

TABLE OF MEANS

.49115	.46869	.47992
.44275	.40558	.42416
.47413	.49778	.48595
.46934	.45735	.46335

VARIABLE 5 ORDERED					
VARIABLE 4 ORDERED	LINEAR	QUADRATIC	CUBIC	REMAINDER	TOTAL
LINEAR	.03083	0	0	0	.03083
QUADRATIC	0	0	0	0	0
CUBIC	0	0	0	0	0
REMAINDER	0	0	0	0	0
TOTAL	.03083	0	0	0	.03083

TABLE OF INTERACTIONS FOR
VARIABLE 4 (ROW)
VARIABLE 5 (COLUMN)

-.01690	.01690	-.00063
.01690	-.01690	.00063
.00600	-.00600	.46335

TABLE OF MEANS

.45182	.47361	.46271
.48687	.44109	.46398
.46934	.45735	.46335

PROBLEM NO. 2

CELL NUMBERS					MEANS
1	1	1	1	2	.32175
1	1	1	2	1	.46365
1	1	1	2	2	.32175
1	1	2	1	1	.32175
1	1	2	1	2	.36137
1	1	2	2	1	.32175
1	1	2	2	2	.32175
1	1	3	1	1	.46365
1	1	3	1	2	.32175
1	1	3	2	1	.32175
1	1	3	2	2	.52360
1	2	1	1	1	.46365
1	2	1	1	2	.46365
1	2	1	2	1	.88608
1	2	1	2	2	.32175
1	2	2	1	1	.32175
1	2	2	1	2	.46365
1	2	2	2	1	.46365
1	2	2	2	2	.32175
1	2	3	1	1	.36137
1	2	3	1	2	.46365
1	2	3	2	1	.32175
1	2	3	2	2	.36137
1	3	1	1	1	.46365
1	3	1	1	2	.32175
1	3	1	2	1	.88608
1	3	1	2	2	.32175
1	3	2	1	1	.10715
1	3	2	1	2	.46365
1	3	2	2	1	.68472
1	3	2	2	2	.29284
1	3	3	1	1	.32175
1	3	3	1	2	.32175
1	3	3	2	1	.32175
1	3	3	2	2	.32175
2	1	1	1	1	.32175
2	1	1	1	2	.24905
2	1	1	2	1	.46365
2	1	1	2	2	.46365
2	1	2	1	1	.46365
2	1	2	1	2	.32175
2	1	2	2	1	.88608
2	1	2	2	2	.46365
2	1	3	1	1	.32175
2	1	3	1	2	.32175
2	1	3	2	1	.32175
2	1	3	2	2	.32175
2	2	1	1	1	.46365
2	2	1	1	2	.46365
2	2	1	2	1	.32175
2	2	1	2	2	.46365
2	2	2	1	1	.32175
2	2	2	1	2	.46365
2	2	2	2	1	.32175
2	2	2	2	2	.46365

2	2	2	2	2	.32175
2	2	3	1	1	.68472
2	2	3	1	2	.46365
2	2	3	2	1	.32175
2	2	3	2	2	.46365
2	3	1	1	1	.32175
2	3	1	1	2	.32175
2	3	1	2	1	.46365
2	3	1	2	2	.46365
2	3	2	1	1	.32175
2	3	2	1	2	.32175
2	3	2	2	1	.36137
2	3	2	2	2	.32175
2	3	3	1	1	.52368
2	3	3	1	2	.68472
2	3	3	2	1	.32175
2	3	3	2	2	.46365
3	1	1	1	1	.46365
3	1	1	1	2	.78540
3	1	1	2	1	.32175
3	1	1	2	2	.68472
3	1	2	1	1	.32175
3	1	2	1	2	.46365
3	1	2	2	1	.32175
3	1	2	2	2	.46365
3	1	3	1	1	.46365
3	1	3	1	2	.32175
3	1	3	2	1	.24905
3	1	3	2	2	.46365
3	2	1	1	1	.88609
3	2	1	1	2	.36137
3	2	1	2	1	.68472
3	2	1	2	2	.32175
3	2	2	1	1	.32175
3	2	2	1	2	.46365
3	2	2	2	1	.46365
3	2	2	2	2	.68472
3	2	3	1	1	.88609
3	2	3	1	2	.46365
3	2	3	2	1	.68472
3	2	3	2	2	.68472
3	3	1	1	1	.32175
3	3	1	1	2	.32175
3	3	1	2	1	.46365
3	3	1	2	2	.32175
3	3	2	1	1	.32175
3	3	2	1	2	.46365
3	3	2	2	1	.32175
3	3	2	2	2	.32175
3	3	3	1	1	.32175
3	3	3	1	2	.88609
3	3	3	2	1	.32175
3	3	3	2	2	.110715

M A R G I N A L M E A N S				
VARIABLES	C A T E G O R I E S			M E A N S
1	1			.42858
	2			.44130

	3	.52016
2	1	.46197
	2	.47767
	3	.45040
3	1	.47992
	2	.42414
	3	.48595
4	1	.46271
	2	.46398
5	1	.46934
	2	.45735

TWAOV

Donald A. Gardiner

(Algol)

DESCRIPTION

TWAOV is a calling program for the procedure THREEWAYANOVA. The procedure will perform an analysis of variance on a $T \times R \times S$ table with U observations per cell. The sums of squares are computed by adding squared deviations from means rather than from equivalent short formulas. That is, $\sum (x_{ij} - \bar{x})^2$ is used rather than $\sum x_{ij}^2 - (\sum x_{ij})^2/N$. The total sum of squares (with $RSTU$ degrees of freedom) is calculated twice: once by adding the squared observations and again by adding the sums of squares of the effects. The second sum is output as a check total and is independent of the first sum because of the manner in which the sums of squares for the effects are calculated.

The input includes the number of levels for the first, second and third classification and the number of observations per cell and the data from the experiment.

The output includes the Tables of Means and the Analysis of Variance Tables. Under the Tables of Means three tables of means are printed. The first is called First Cross by Second Cross and is a table of averages over the third classification. It is augmented by a row of column means and a column of row means. The second table is called First Cross by Third Cross and the third table is called Second Cross by Third Cross. They have similar content.

The analysis of variance table consists of two parts if the number of levels of the first classification is greater than one, or of only one part if the number of levels of the first classification is equal to one. The first part of the analysis of variance table is an analysis of a three way cross classification with equal numbers of observations per cell. The second part is an analysis of a nested design in four factors.

For the three way cross classification F ratios are computed using as the denominator the mean square of the first of the following effects which has positive degrees of freedom:

- Error (within cells)
- Three factor interaction
- Second by third interaction
- First by third interaction
- First by second interaction.

All F ratios have the same denominator in the cross classification analysis.

In the hierarchal analysis, the F ratios are computed in the proper manner, there being only one proper manner in the absence of pooling.

CARD PREPARATION

Order of Cards

1. Master Control System Card
2. Δ Execute Card
3. Binary Program Deck
4. Δ Card

- 5. Title Card
- 6. Levels Card
- 7. Format Card
- 8. Data Cards

Items 5-8 may be repeated as many times as desired.

Master Control System Card

Starting in Column 1 the following characters appear:

ΔALGOL, Request Number, Name.

Title Card

Any or all of the eighty columns on this card may be used for alphanumeric problem identification. This card must be included even if blank.

Levels Card

<u>Columns</u>	<u>Explanation</u>
1-3	Number of levels of the first classification (T)
4-6	Number of levels of the second classification (R)
7-9	Number of levels of the third classification (S)
10-12	Number of observations per cell (U)

Format Card

Columns 1-80 may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats. A format for the data must be specified.

Data Cards

The data cards should be arranged in the following sequence where T corresponds to the 1st cross, R corresponds to the 2nd cross, S corresponds to the 3rd cross, and U corresponds to repeated observations.

For example, if $T = 2$, $R = 2$, $S = 2$, $U = 2$ the data would be arranged according to the following sequence.

<u>T</u>	<u>R</u>	<u>S</u>	<u>U</u>
1	1	1	1
1	1	1	2
2	1	1	1
2	1	1	2
1	2	1	1
1	2	1	2
2	2	1	1
2	2	1	2
1	1	2	1
1	1	2	2
2	1	2	1
2	1	2	2
1	2	2	1
1	2	2	2
2	2	2	1
2	2	2	2

EXAMPLE

In the example for TWAOV $T = 4$, $R = 3$, $S = 2$, and $U = 2$. There is a single observation per card.

TEST DATA FOR TWAEV
4 3 2 2

(F10,0)

41.2
42.7
42.6
42.9
44.3
46.0
45.0
48.4
40.1
41.2
40.8
41.6
39.4
41.0
42.3
39.0
36.4
34.1
40.2
38.6
41.8
43.2
45.4
47.1
54.4
53.3
57.6
55.4
40.3
43.7
56.6
59.2
47.5
45.7
46.1
48.7
41.3
39.0
42.4
46.1
36.8
34.0
38.1
41.0
39.7
42.6
39.0
40.9

TEST DATA FOR TWOV
TABLES OF MEANS

FIRST CROSS BY SECOND CROSS

ROW	1	2	3	ROW MEAN
1	4.7900000+01	4.3625000+01	3.5325000+01	4.2283333+01
2	4.9625000+01	4.4300000+01	3.9475000+01	4.4466667+01
3	5.3575000+01	4.0175000+01	4.1825000+01	4.5191667+01
4	5.2300000+01	4.2450000+01	4.3100000+01	4.5950000+01
COLUMN MEANS	5.0850000+01	4.2637500+01	3.9931250+01	4.4472917+01

FIRST CROSS BY THIRD CROSS

ROW	1	2	ROW MEAN
1	3.9283333+01	4.5283333+01	4.2283333+01
2	4.1166667+01	4.7816667+01	4.4466667+01
3	4.2616667+01	4.7766667+01	4.5191667+01
4	4.4533333+01	4.7366667+01	4.5950000+01
COLUMN MEANS	4.1887500+01	4.7058333+01	4.4472917+01

SECOND CROSS BY THIRD CROSS

ROW	1	2	ROW MEAN
1	4.4137500+01	5.7562500+01	5.0850000+01
2	4.0675000+01	4.4600000+01	4.2637500+01
3	4.0850000+01	3.9012500+01	3.9931250+01
COLUMN MEANS	4.1887500+01	4.7058333+01	4.4472917+01

ANALYSIS OF VARIANCE

SOURCE	D.F.	SUMS OF SQUARES	MEAN SQUARES	F RATIOS
TOTAL	48	9,7161030E+04		
MEAN	1	9,4936335E+04		
MAIN EFFECTS				
FIRST	3	8,9912292E+01	2,9970764E+01	1,1380402E+01
SECOND	2	1,0346029E+03	5,1730146E+02	1,9642805E+02
THIRD	1	3,2085021E+02	3,2085021E+02	1,2183221E+02
INTERACTIONS				
FIRST BY SECOND	6	1,6857708E+02	2,8096181E+01	1,0668592E+01
FIRST BY THIRD	3	2,5470625E+01	8,4902083E+00	3,2238747E+00
SECOND BY THIRD	2	4,7520042E+02	2,3760021E+02	9,0220789E+01
THREE FACTOR	6	4,6876250E+01	7,8127083E+00	2,9666166E+00
ERROR	24	6,3205000E+01	2,6335417E+00	1,0000000E+00
CHECK TOTAL	48	9,7161030E+04		

HIERARCHAL ANALYSIS

FIRST	3	8,9912292E+01	2,9970764E+01	1,9927701E+01
SECOND IN FIRST	8	1,2031800E+03	1,5039750E+02	2,0782764E+00
THIRD IN SECOND	12	8,6839750E+02	7,2366458E+01	2,7478760E+01
ERROR	24	6,3205000E+01	2,6335417E+00	1,0000000E+00

YATES

Norma C. Hull

(Fortran 63)

DESCRIPTION

This program performs an analysis of a 2^N factorial experiment by the method known as the "Yates calculation." Several problems may be run, one behind another. The number of factors is limited to 7, that is, $N \leq 7$. The program can handle 2^N factorial designs with or without multiple observations.

The input includes the number of factors, the number of observations per treatments, alphanumeric identification, and the observations from the experiment.

The output includes:

- (1.) An analysis of variance table,
- (2.) A table of five columns containing the pertinent information of the Yates calculation:
 - (a.) a column identifying the effects,
 - (b.) a column of the means of the responses for the 2^N factor combinations,
 - (c.) a column which is the result of performing the successive additions and successive subtractions in the Yates method N times,
 - (d.) a column of estimates of the effects,
 - (e.) a column containing the sum of squares attributable to each effect,
- (3.) The sum of squares of deviations from the experiment mean,
- (4.) A graph of the Cuthbert Daniel half-normal plot.

CARD PREPARATION

Order of Cards

1. Master Control System Card
2. Δ Execute Card
3. Binary Program Deck
4. Δ Card
5. Problem Card
6. Factor Name Cards
7. Format Card(s)
8. Data

Master Control System Card

Starting in column 1 the following characters appear:

Δ COOP, REQUEST NUMBER, NAME, ϕ /HD2, G.

Problem Card

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	PR	PRoblem
3	N	Number of factors $N \leq 7$
4,5	M	Number of observations per treatment $m < 100$
6,7	NF	Number of format cards, $1 \leq NF \leq 5$
8,71		Alphanumeric problem identification

Factor Name Cards

These cards contain any alphanumeric information to identify the factors of the experiment. The number of cards must equal N.

Columns

1,72

Alphanumeric identification

Format Card(s)

Columns 1-72 of up to five cards may be used to describe the data using E or F conversions. See Appendix C for details concerning formats. A format for the data must be specified.

Data Card(s)

Data must be punched in agreement with the format cards. The data must be arranged in standard factorial order, i.e., (1) a, b, ab, c, ac, bc, abc, ...,

EXAMPLES

The two examples for YATES are similar. The exception being that in the first example there are 3 replications and in the second example there is only one observation per point. The plots for each example present a graphic picture of the Effects Column.

The 1604A job card must show that tape drive "2" is used for plotting.

PR4 3 IPROBLEM TC TEST PROGRAM YATES
ACID STRENGTH
TIME OF REACTION
AMOUNT OF ACID
TEMPERATURE OF REACTION
(F4,2,2X,F4,2,2X,F4,2)

,05	0,08	0,11
,01	0,04	0,07
,50	0,53	0,56
,40	0,43	0,46
,28	0,31	0,34
,06	0,09	0,12
,09	0,12	0,15
,33	0,36	0,39
,76	0,79	0,82
,65	0,68	0,71
,70	0,73	0,76
,05	0,08	0,11
,74	0,77	0,80
,35	0,38	0,41
,46	0,49	0,52
,20	0,23	0,26

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

YATES CALCULATION ON A 2 TO THE N FACTORIAL

PROBLEM TO TEST PROGRAM YATES

TREATMENT A= ACID STRENGTH

TREATMENT B= TIME OF REACTION

TREATMENT C= AMOUNT OF ACID

TREATMENT D= TEMPERATURE OF REACTION

COMBINATION	SUM X(I)/M	NTM COLUMN	EFFECT	SUM OF SQUARES
I	8,000000000**02	6,110000000**00	3,818750000**01	6,999768750**00
A	4,000000000**02	-1,530000000**00	-1,912500000**01	4,389187500**01
B	5,300000000**01	-1,699999999**01	-2,124999999**02	5,418749995**03
AB	4,300000000**01	-9,999999995**03	-1,249999999**03	1,874999998**05
C	3,100000000**01	-6,100000000**01	-7,624999999**02	6,976874999**02
AC	9,000000000**02	2,700000000**01	3,375000000**02	1,366875000**02
BC	1,200000000**01	-5,300000000**01	-6,625000000**02	5,266875000**02
ABC	3,600000000**01	1,196000000**00	1,487500000**01	2,655187500**01
D	7,900000000**01	2,190000000**00	2,737500000**01	8,992687499**01
AD	6,800000000**01	-1,290000000**00	-1,612500000**01	3,120187500**01
BD	7,300000000**01	-2,010000000**00	-2,512500000**01	7,575187500**01
ABD	8,000000000**02	-8,100000000**01	-1,012500000**01	1,230187500**01
CD	7,700000000**01	-2,100000000**01	-2,625000000**02	8,268749999**03
ACD	3,800000000**01	-4,999999998**02	-6,249999998**03	4,687499997**04
BCD	4,900000000**01	9,900000000**01	1,237500000**01	1,837687500**01
ABCD	2,300000000**01	1,500000000**01	1,875000000**02	4,218750001**03
TOTAL	6,110000000**00			1,013430000**01

CORRECTED SUM OF SQUARES OF THE 48 RESULTS= 3,134531250**00

ANALYSIS OF VARIANCE OF DATA

SOURCE	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARE
TOTAL	48	1,016310000 $\times 10^1$	
MEAN	1	6,999768750 $\times 10^0$	
EFFECTS	15	3,134531250 $\times 10^0$	2,10887500 $\times 10^1$
REMAINDER	32	2,880000044 $\times 10^2$	9,000000136 $\times 10^4$

PROBLEM TO TEST PROGRAM YATES

PR4 1 IPRBLEM NC,2 TO TEST YATES
ACID STRENGTH
TIME OF REACTION
AMOUNT OF ACID
TEMPERATURE OF REACTION
(F4,2)

,08

,04

,53

,43

,31

,09

,12

,36

,79

,68

,73

,08

,77

,38

,49

,23

STATISTICS SECTION OF MATHEMATICS DIVISION OF OAK RIDGE NATIONAL LABORATORY

YATES CALCULATION ON A 2 TO THE N FACTORIAL

PROBLEM NO. 2 TO TEST YATES

TREATMENT A= ACID STRENGTH

TREATMENT B= TIME OF REACTION

TREATMENT C= AMOUNT OF ACID

TREATMENT D= TEMPERATURE OF REACTION

COMBINATION	SUM X(I)/M	ITH COLUMN	EFFECT	SUM OF SQUARES
I	8,000000000**02	6,110000000**00	-3,818750000**01	2,333256250**00
A	4,000000000**02	-1,530000000**00	-1,912500000**01	1,463062500**01
B	5,300000000**01	-1,700000000**01	-2,124999999**02	1,806249999**03
AB	4,300000000**01	-9,999999987**03	-1,249999998**03	6,249999984**06
C	3,100000000**01	-6,100000000**01	-7,624999999**02	2,325625000**02
AC	9,000000000**02	2,700000000**01	3,375000000**02	4,556250000**03
BC	1,200000000**01	-5,300000000**01	-6,625000000**02	1,755625000**02
ABC	3,600000000**01	1,190000000**00	1,487500000**01	8,850624999**02
D	7,900000000**01	2,190000000**00	2,737500000**01	2,997562500**01
AD	6,800000000**01	-1,290000000**00	-1,612500000**01	1,040062500**01
BD	7,300000000**01	-2,010000000**00	-2,512500000**01	2,525062500**01
ABD	8,000000000**02	-8,100000000**01	-1,012500000**01	4,100625000**02
CD	7,700000000**01	-2,100000000**01	-2,625000000**02	2,756250000**03
ACD	3,800000000**01	-4,999999999**02	-6,249999999**03	1,562500000**04
BCD	4,900000000**01	9,900000000**01	1,237500000**01	6,125625000**02
ABCD	2,300000000**01	1,500000000**01	1,875000000**02	1,406250000**03
TOTAL	6,110000000**00			3,378100000**00

CORRECTED SUM OF SQUARES OF THE 16 RESULTS= 1,044843750**00

ANALYSIS OF VARIANCE OF DATA

SOURCE	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARE
TOTAL	16	3,378,00000+00	
MEAN	1	2,333256250+00	
EFFECTS	15	1,044843750+00	6,965625000+02
REMAINDER	0	5,820766091+11	0

PROBLEM NO. 2 TO TEST YATES

GRAPH

D. A. Gardiner

(ALGOL)

DESCRIPTION

The program, GRAPH, will produce a graph of bivariate data and compute elementary correlation statistics. The bivariate data are plotted as unconnected points by the Calcomp plotter. The printed output consists of a list of the data, the means and standard deviations of each variable, the correlation coefficient, the t statistic under the hypothesis of zero correlation, the 0.975 quantile of Student's t distribution and a statement that the null hypothesis is either rejected or not rejected at the five percent level of percent level of significance.

CARD PREPARATION

Order of Cards

1. Master Control System Card
2. Δ EXECUTE Card
3. Program Deck
4. Title Card
5. Sample Size Card
6. Format Card
7. Data Deck

Items 4, 5, 6, and 7 may be repeated as often as desired.

Master Control System Card

Starting in column 1 the following characters appear:

ΔALGOL, Request number, name, Ø17.

Title Card

Any alphanumeric information may be punched in any or all of the columns of the title card. Whatever is punched in this card will be reproduced at the top of the graph and on the first line of the printed output.

We recommend that the title identify the variables since the labels on the graph are merely "Y-AXIS" and "X-AXIS."

Sample Size Card

In columns 1, 2, and 3 punch the number of data points (pairs of observations) in the sample. The last digit of this number should be punched in column 3. Blanks are interpreted as zeroes.

Format Card

Columns 1-72 may be used to describe the input data using E or F conversions. See Appendix C for details concerning formats. A format for the data must be specified.

Data Deck

The program reads two numbers at a time. The first number is taken to be the coordinate of the X-variable of a point and the second number is taken to be the coordinate of the Y-variable for the same point.

EXAMPLE

The example for GRAPH has 22 values of an independent variable and dependent variable. The output indicates that there is a correlation between the two variables.

The 1604A job card must indicate that tape drive "17" is used for plotting. The Calcomp plotting card must show $Y = 0$ at $1/2$ ".

EXPERIMENT NO. 6 X IS LOG(SR/ZR), Y IS LOG(BA/ZR) FULL SET OF DATA

22
 (2F10,0)
 1.837543 .748335
 2.352612 2.008450
 1.907147 2.006356
 12.428148 10.173795
 4.574147 3.779658
 1.356406 2.489454
 6.520905 4.845733
 1.780927 1.826211
 10.009843 7.928739
 1.896044 3.100981
 3.655769 3.211282
 2.337406 2.949299
 2.421706 3.058857
 .086828 .042799
 .349318 .157967
 .031861 .014872
 .043809 .022483
 .009925 .006671
 .003825 .002340
 .002364 .001237
 .180095 .115618
 .002712 .002375

EXPERIMENT NO. 6 X IS LOG(SR/ZR), Y IS LOG(BA/ZR) FULL SET OF DATA

	X-VARIABLE	Y-VARIABLE
	1.837543	.748335
	2.352672	2.008450
	1.907147	2.006396
	12.428148	10.173795
	4.574147	3.779658
	1.356486	2.489454
	6.520905	4.845733
	1.780927	1.826211
	10.009843	7.928739
	1.896044	3.100981
	3.655769	3.211282
	2.337406	2.949299
	2.421706	3.058857
	.086828	.042799
	.349318	.157967
	.031861	.014872
	.043809	.022483
	.009925	.006671
	.003825	.002340
	.002364	.001237
	.180095	.115618
	-.002712	.002375
MEAN	2.444970	2.204252
STANDARD DEVIATION	3.328094	2.702928
CORRELATION COEFFICIENT		.977656
T STATISTIC		20.798963
STUDENTS T		2.086224
HYPOTHESIS OF ZERO CORRELATION IS REJECTED AT THE FIVE PERCENT LEVEL		

EXPERIMENT NO.6 X IS LOG (SR/ZR), Y IS LOG (BR/ZR) FULL SET OF DATA

ELLIPSE

Norma C. Hull

(Fortran 63)

DESCRIPTION

This program plots the simultaneous confidence ellipse about the intercept, A, and the slope, B, for the simple linear regression model, $Y = A + BX$. Using the expression

$$\Pr \left\{ \frac{k(a - A - B\bar{X}) + (b - B)^2 \sum (x_i - \bar{x})^2}{2s^2} < F_{1-\alpha}(2, k-2) \right\} = 1 - \alpha$$

where

k = number of observations

a = $\bar{y}$

b = least squares estimate of B

$s^2 = [\sum_{i=1}^k (y_i - \bar{y})^2] / k - 2$, all pairs of values of (A, B) which satisfy this inequality lie inside the 100 (1- α) per cent confidence region.¹

The boundary of the region is calculated by substituting an equality symbol for the inequality symbol and solving for A. If $F = F_{1-\alpha}(2, k-2)$, then

$$A = (\bar{y} - B\bar{x}) \pm \left\{ \frac{1}{k} \left\{ 2s^2 F - (b - B)^2 \sum (x_i - \bar{x})^2 \right\}^{1/2} \right\}.$$

Two levels of significance may be used which generate concentric ellipses. A slight modification of the program permits confidence

¹Brownlee, K. A., Statistical Theory and Methodology in Science and Engineering. New York: John Wiley and Sons, Inc., 1960.

ellipses of two sets of data to be plotted with the ellipses of like F values of the two data sets compared on 2 graphs (ELLIPSE VERSION 2). The input information may be readily obtained from either BMD29GR or BMD31GR mentioned in this memo. The input consists of

- (1.) Number of observations,
- (2.) Mean value of X, Y,
- (3.) Error mean square,
- (4.) Sum of squares of deviations of x from $\bar{x}$,
- (5.) Least squares estimate of B,
- (6.) F-values.

The output is the graph of the concentric ellipse or ellipses.

CARD PREPARATION

Order of Cards

- (1.) Master Control System Card
- (2.) Δ EXECUTE Card
- (3.) Binary Program Deck
- (4.) Δ Card
- (5.) Number of Iterations for Each Ellipse Side
- (6.) Input Data Cards

As many Data Cards as desired may be included for each run. Using ELLIPSE each Data Card produces a graph; and using ELLIPSE VERSION 2, each pair of Data Cards produces two graphs.

Master Control System Card

Starting in Column 1 the following characters appear:

Δ COOP, Request Number, Name, ϕ /HD2, G.

Number of Iterations Card

In Columns 1-3 specify the number of iterations desired to compute each side of the ellipse. The examples were drawn with 150 iterations per side.

Input Data Cards

<u>Columns</u>	<u>Variable</u>	<u>Explanation</u>
1-5	VK	The number of observations (k)
6-15	YB	The mean value of Y ($\bar{Y}$)
16-25	XB	The mean value of X ($\bar{X}$)
26-35	SSQ	The error mean square (s^2) expressed with as many significant digets as desired i.e. X.XXXXX
36-38	IXPS	The exponent which indicates the power of ten by which the error mean square is to be multiplied by. This is an integer variable and must be right justified.
39-48	XXMBS	Sum of squares of deviations of x from $\bar{x}$ expressed with significant digets ($\sum(x_i - \bar{x})^2$)
49-51	IXPX	The exponent associated with the sum of squares. Must be a right justified integer.
52-61	B	The least squares estimate of B expressed with significant digets.
62-64	IXPB	The exponent associated with the estimate of B . Must be a right justified integer.

65-69	F1	First F value $F(\alpha = 0.01)$
75-79	F2	Second F value $F(\alpha = 0.10)$

EXAMPLES

The two examples for ELLIPSE and ELLIPSE VERSION 2 are the three following graphs. The first graph is produced by ELLIPSE and the second and third graphs are produced by ELLIPSE VERSION 2.

The 1604A job card must show that tape drive "2" is used for plotting.

150
27. 1.9101 803.48 3.1844 -4 4.5052 5-4.8814 -4 5.57 3.38 2.53

159

27	1.8581	803.48	2.9564	-4 4.5052	5 5.3295	-4 5.57	2.53
27	1.9101	803.48	3.1844	-4 4.5052	5 5.4814	-4 5.57	2.53

Appendix A
Transgeneration List for BMD29GR and BMD31GR

The term transgeneration is used to include transformations of input variables and creation of new variables prior to the normal computations performed by the various programs.

TRANSGENERATION LIST

<u>Code</u>	<u>Transformation</u>	<u>Restriction</u>
01	$X' = \sqrt{X}$	$X \geq 0$
02	$X' = \sqrt{X} + \sqrt{X} + 1$	$X \geq 0$
03	$X' = \text{LN} X$	$X > 0$
04	$X' = e^X$	----
05	$X' = \text{SIN}^{-1} \sqrt{X}$	$0 \leq X \leq 1$
06	$X' = \text{SIN}^{-1} \frac{X}{N+L} + \text{SIN}^{-1} \frac{X+1}{N+1}$	$0 \leq \frac{X}{N} \leq 1$
07	$X' = 1/X$	$X \neq 0$
08	$X' = X+C$	----
09	$X' = X * C$	----
10	$X' = X^C$	$X > 0$
11	$X' = X_A + X_B$	----
12	$X' = X_A - X_B$	----
13	$X' = X_A * X_B$	----
14	$X' = X_A/X_B$	$X \neq 0$

15	X_A Used As Weight	
16	$X' = \sin X$	----
17	$X' = \cos X$	----
18	$X' = \tan X$	----
19	$X' = \sin^{-1} X$	$1 \times 1 \leq 1$
20	$X' = \cos^{-1} X$	$1 \times 1 \leq 1$
21	$X' = \sinh X$	----
22	$X' = \cosh X$	----
23	$X' = (X) $	----
24	$X' = X_A \text{ MODULO } C$	$C \neq 0$
25	$X' = \sqrt[3]{X}$	----
26	$X' = X_A^{X_B}$	$X_A > 0$
27	$X' = C^{X_A}$	$C > 0$

When using Code 15 as weight the Trans-Generation Card should read:

<u>Columns</u>	<u>Entry</u>	<u>Explanation</u>
1,2	TR	<u>T</u> ransgeneration Card
3,4	BLANK	
5,6	15	Transgeneration Code
7,8		A-variable number to be used as the weight

Appendix B

Transgeneration List for ANOVA, ANOVAGR, GENLIN, and NONLIN

The term transgeneration is used to include transformations of input variables and creation of new variables prior to the normal computations performed by the various programs.

Cols. 1,2	3,4	5	6,7	8-15	16,17	18-25	Transformation
TR	x'	=		SQRT	x		$(x') = \sqrt{(x)}$
TR	x'	=		LN	x		$(x') = \ln(x)$
TR	x'	=		EXP	x		$(x') = e^{(x)}$
TR	x'	=		ABS	x		$(x') = (x) $
TR	x'	=		RECIP	x		$(x') = 1/(x)$
TR	x'	=	x_A	PLUS	x_B	C	$x_B \neq 0 \quad (x') = (x_A) + (x_B)$ $x_B = 0 \quad (x') = (x_A) + C$
TR	x'	=	x_A	MINUS	x_B	C	$x_B \neq 0 \quad (x') = (x_A) - (x_B)$ $x_B = 0 \quad (x') = (x_A) - C$
TR	x'	=	x_A	TIMES	x_B	C	$x_B \neq 0 \quad (x') = (x_A) * (x_B)$ $x_B = 0 \quad (x') = (x_A) * C$
TR	x'	=	x_A	OVER	x_B	C	$x_B \neq 0 \quad (x') = (x_A) / (x_B)$ $x_B = 0 \quad (x') = (x_A) / C$
TR	x'	=	x_A	TOTHE	x_B	C	$x_B \neq 0 \quad (x') = (x_A)^{(x_B)}$ $x_B = 0 \quad (x') = (x_A)^C$
TR	x'	=		SIN	x		$(x') = \sin(x)$
TR	x'	=		COS	x		$(x') = \cos(x)$
TR	x'	=		ARCSIN	x		$(x') = \sin^{-1}(x)$
TR	x'	=		ARCTAN	x		$(x') = \tan^{-1}(x)$
TR	x'	=		ARCSINH	x		$(x') = \sinh(x)$
TR	x'	=		ARCCOSH	x		$(x') = \cosh(x)$
TR	x'	=		PROBIT	x		$(x) = \int_{-\infty}^{(x')-5} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$
TR	x	=		WEIGHT			

Remarks

- (1.) Columns 8-15 must be left justified.
- (2.) (x) is to be read "contents of x."
- (3.) (x), (x_A), (x_B), and C are unchanged by a transgeneration.
- (4.) Up to 99 transgeneration cards may be used for a problem.
- (5.) C is read by an F8.0 Format. A punched decimal point will override the decimal point assumed to the right of column 25.

Appendix C

The word "format" usually refers to the arrangement of information keypunched on a card. The format of a data card is a sequence of fields (variables), each of which occupies one or more columns. For the computer programs a format is a set of specifications according to which information is read into the program from punched cards. The specifications tell the program which parts (or columns) of the card to skip, which parts to regard as all one number, and which parts to regard as several numbers in a row. For instance, it is the format which tells the program whether a card punched "345890021" is to be read in as "34.5, 890.0, .021", or "34., 9002.1", or "589.002", or "3, 4, 58, 90, 0, 2, 1", etc. It does this by giving the program a sequence of specifications which indicate the size of a field and the method of handling that field (i.e., skipping it, entering it into the computer as a whole number, entering it into the computer as a number with two decimal digits, etc.).

The format also tells the program how to read in a certain set of cards when more than one are required to contain all the data for one case. In most programs the format describes the variables for each case. Each successive case is assumed to have the same format.

If the formats for the programs were fixed in advance, all data would have to be punched on cards in the same way for every study. Since this is not usually convenient, the programs have been written so that the user may vary the formats according to his preference for a particular study. For this reason, they are referred to as

"variable formats". The program is informed of the format which is to be used by Variable Format Cards.

In addition to providing an economical method of preparing data input cards (by defining fields to be as small as possible, or "packing" the data), the variable format permits considerable freedom in controlling data input. For instance:

It allows the user to select for each case only those cards which have fields of interest. (See Examples ii, iii, vi below.)

It allows the user to select only those fields of interest from among the fields of each card. (See Examples iv, v, vi.)

It allows the user to scale the data input, i.e., shift the decimal point.

1. F-type Variable Format

The F-type format is the most frequently used in the programs. It is required when the decimal point is keypunched on the card or when the decimal point is to be placed by the program. All data input values must be signed (+) or unsigned numbers with or without a decimal point punched.

Specifications:

- (a) "nFw.d" F is the floating point (decimal) indicator;
n is the number of fields of width w (includes sign and decimal point if punched); and d is the number of digits to the right of the decimal point if the decimal is not punched ($0 \leq d \leq w$). If the decimal is punched, d is ignored. If n is not specified, it is assumed to be 1.
- (b) "sPnFw.d" P is the scale indicator;
s is a scale factor (explained below); and n, w and d are defined in specification (a).
- (c) "mX" (alphabetic X) X is the skip indicator, and m is the number of columns to be skipped.
- (d) "/" (slash) indicates "go to the next data card."

Depending on its location in the format statement, the "/" will either direct the program to go immediately to the next card (ignoring any further information on the current card) or skip one card altogether. For example,

if a format begins with "/", the program will automatically skip the first card, read the second, skip the third card, etc. If a format ends with "/", the program will automatically read the first card, skip the second, read the third card, etc. (See Examples ii, iii, v, vi.)

"//" indicates "go to the card after next."

Two slashes "/" will direct the program to skip two cards. Any number of slashes may be used.

The format is keypunched beginning with a left parenthesis, a sequence of specifications, and closed by a right parenthesis. Specifications (a), (b), and (c) are followed by a comma, except preceding a slash or right-hand parenthesis. Blank columns within the format are ignored.

Examples:

- (i) (12F3.0, F4.0, 11F2.0), punched in the first 20 columns of the Variable Format Card, will describe 12 three-column fields followed by 1 four-column field, followed by 11 two-column fields. Each data card will be read according to this format.
- (ii) (12F3.0, F4.0, 11F2.0 /), punched in the first 22 columns, will describe the same fields as in Example (i), but will also instruct the program to read the first card, skip the second, read the third card, skip the fourth, etc.

Two slashes at the end

..., 11F2.0 //)

will instruct the program to read the first card, skip the next two cards, read the fourth card, skip the next two, read the seventh card, etc.

- (iii) (/12F3.0, F4.0, 11F2.0), punched in the first 21 columns, will describe the same fields as in Example (i), but will also instruct the program to skip the first card, read the second, skip the third card, read the fourth, etc.

Two slashes at the beginning

(//12F3.0,...

will instruct the program to skip the first two cards, read the third, skip the next two, read the sixth card, skip the next two, etc.

(iv) (10X,F6.0,2X,2F3.0), punched in the first 19 columns, will instruct the program to direct entry of data from each card as follows:

- (1) Skip 10 columns.
- (2) Pick up a six-digit field in Col. 11-16.
- (3) Skip 2 columns.
- (4) Pick up 2 three-digit fields in Col. 19-21 and 22-24.

(v) (5X,2F6.0,F1.0,3X,F5.0 /5X,F6.0), punched in the first 32 columns, will instruct the program to direct entry of data from each pair of cards as follows:

- (1) Skip 5 columns.
- (2) Pick up 2 six-digit fields in Col. 6-11 and 12-17.
- (3) Pick up a one-digit field in Col. 18.
- (4) Skip three columns (Col. 19-21).
- (5) Pick up a five-digit field in Col. 22-26.
- (6) Go to second data card.
- (7) Skip 5 columns.
- (8) Pick up a six-digit field in Col. 6-11 of second card.
- (9) Repeat for each pair of data cards.

(vi) (5X,2F6.0,F1.0,3X,F5.0 //5X,F6.0,/F4.0,2X,F1.0), punched in the first 47 columns, will instruct the program to direct entry of data from each set of four cards as follows:

- (1)-(5) Same as Example v.
- (6) Skip second card and go to third card.
- (7) Skip 5 columns.
- (8) Pick up a six-digit field in Col. 6-11 of third card.
- (9) Go to fourth card.
- (10) Pick up a four-digit field in Col. 1-4 of fourth card.
- (11) Skip 2 columns.
- (12) Pick up a one-digit field in Col. 7.
- (13) Repeat for each set of four cards.

SCALING

Scaling may be indicated by using either the "nFw.d" specification of the "sPnFw.d" specification. This specification will not often be required.

"nFw.d" Specification:

When the decimal point is not punched, the d of the above specification instructs the program to divide the whole number picked up by 10^d . For example, F6.1 will specify that the number picked up in a six-column field be divided by 10, F6.2 will specify division by 100, and F6.6 will specify division by 1000000.

$$\text{Stored number} = \text{punched number} / 10^d.$$

When the decimal point is punched, d is ignored.

Examples:

<u>Punched Number</u>	<u>Format Specification</u>	<u>Stored Number</u>
2468	F4.0	2468.0
3691	F4.1	369.1
4810	F4.3	4.810
4911.32	F7.0	4911.32
4911.32	F7.4	4911.32
172115	F6.6	0.172115

The entire format for these specifications might be punched as:

(F4.0, F4.1, F4.3, F7.0, F7.4, F6.6)

"sPnFw.d" Specification:

Whether the decimal point is punched or not, the s of the above specification instructs the program to divide the number picked up by 10^s . For example, 2PF6.0 will specify division by 100, -2PF6.0 will specify division by .01.

Decimal point punched,

$$\text{Stored number} = \text{punched number} / 10^s, \quad -8 < s < 8$$

Decimal point not punched,

$$\text{Stored number} = \text{punched number} / 10^{s+d}$$

Examples:

<u>Punched Number</u>	<u>Format Specification</u>	<u>Stored Number</u>
7432	1PF4.0	743.2
74.32	-5PF5.0	7432000.0
7432	2PF4.3	.07432
7432	1P2F2.0	7.4 and 3.2
7432	2P2F2.1	0.074 and 0.032
7432	OPF4.1	743.2
7.432	-3PF5.0	7432.0

The entire format for these specifications might be punched as:

(1PF4.0, -5PF5.0, 2PF4.3, 1P2F2.0, 2P2F2.1, OPF4.1, -3PF5.0)

Mixed "nFw.d" and "sPnFw.d" Specifications:

Once the sPnFw.d specification has been used, it will hold for all Fw.d specifications to the right of it until another sPnFw.d is encountered.

If the sPnFw.d specification is not necessary to the right of its occurrence within the variable format statement, then OPFw.d should be used for the next specification to the right, which will hold then for the remaining specifications.

Examples:

(vii) (2F3.1, 2X, F4.2, -6PF3.0, F4.0)

The program interprets the last field as -6PF4.0.

(viii) (2F3.1, 2X, F4.2, -6PF3.0, OPF3.0)

The program interprets the last field as F3.0.

(ix) (1PF3.0, F2.0, OPF3.1, F4.2, F4.0)

The program interprets the second field as 1PF2.0.

2. I-type Variable Format

This format is required for programs designed to process only integer values. The specification is "nIw", where w is the width of the field (includes sign if punched), and n is the number of fields (assumed to be 1 if not punched). All data must be signed (+) or unsigned integers with no decimal point punched. Examples corresponding to those given for F-type are:

- (i) (12I3,I4,11I2)
- (ii) (12I3,I4,11I2 /)
 ...,11I2 //)
- (iii) (/12I3,I4,11I2)
 (//12I3,...
- (iv) (10X,I6,2X,2I3)
- (v) (5X,2I6,1I,3X,I5 /5X,I6)
- (vi) (5X,2I6,1I,3X,I5 //5X,I6, /I4,2X,1I)

No scaling is permitted with I-type format.

3. A-type Variable Format

This format is required for programs designed to process data with alphabetic, numeric, or special characters, or combinations of these. The specification is "nAw", where w is the width of the field, $1 \leq w \leq 8$, and n is the number of fields (assumed to be 1 if not punched). Each specification of a field results in a computer word consisting of exactly 8 characters. When $w < 8$, the characters are positioned in the left of the field and the remaining characters are filled in with blanks. The following examples illustrate certain rules.

<u>Punched Data</u>	<u>Format Specification</u>	<u>Stored* Data</u>
12.0	A4	12.0bb
AGE	A3	AGEbbb
CANCER	A6	CANCER
\$	A1	\$bbbbbb
X+Y=A	A5	X+Y=Ab

*b indicates blank character

INTERNAL DISTRIBUTION

- 1- 2. Central Research Library
- 3. Document Reference Section
- 4- 8. Laboratory Records
- 9. Laboratory Records - Record Copy
- 10- 24. Division of Technical Information Extension
- 25- 74. Mathematics Division (X-10)
- 75-149. Mathematics Division (Y-12)
- 150. ORNL Patent Office

EXTERNAL DISTRIBUTION

- 151. K. O. Bowman (ORGDP)
- 152. E. E. Johnson (Y-12)
- 153. E. L. Frome (ORAU)
- 154. W. T. Butts, Jr. (UTAEC)
- 155. Charles Thigpen
Statistics Department
University of Tennessee
Knoxville, Tennessee
- 156. Research and Development Div., ORO