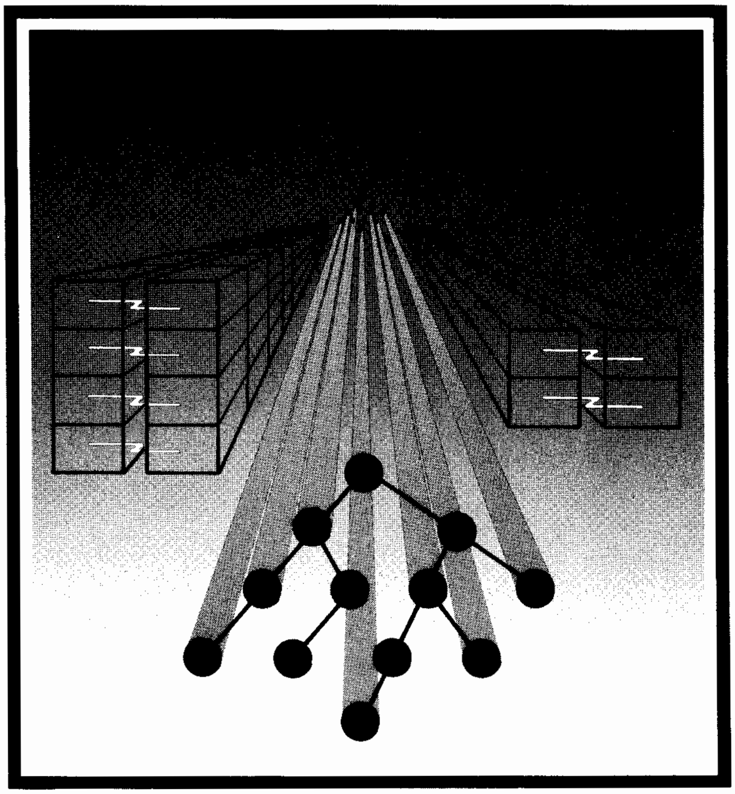
HP 9800 Computer Systems

Analysis of Variance

For the HP 9845



HEWLETT PACKARD

Analysis of Variance

Part No. 09845-15171

Hewlett-Packard Desktop Computer Division
3404 East Harmony Road, Fort Collins, Colorado 80525

Printing History

New editions of this manual will incorporate all material updated since the previous edition. Update packages may be issued between editions and contain replacement and additional pages to be merged into the manual by the user. Each updated page will be indicated by a revision date at the bottom of the page. A vertical bar in the margin indicates the changes on each page. Note that pages which are rearranged due to changes on a previous page are not considered revised.

The manual printing date and part number indicate its current edition. The printing date changes when a new edition is printed. (Minor corrections and updates which are incorporated at reprint do not cause the date to change.) The manual part number changes when extensive technical changes are incorporated.

June 1979...Revision A
August 1980...Revision B. Updated pages: iii, xi, xii, xiii, xiv, 7, 59, 77, 179, 183, 184
May 1981...Revision C. Updated pages: 188
| August 1982...Updated pages 2, 187 thru 190.

Important

The tape cartridge or disc containing the programs is very reliable, but being a mechanical device, is subject to wear over a period of time. To avoid having to purchase a replacement medium, we recommend that you immediately duplicate the contents of the tape onto a permanent backup tape or disc. You should also keep backup copies of your important programs and data on a separate medium to minimize the risk of permanent loss.

Table of Contents

Analysis of Variance

Commentary

The procedure which statisticians call the analysis of variance (ANOVA) has been used for decades as a means of separating the total variability of a set of numbers about the overall average into component parts. The real significance of this separation is that we are able to assign specific sources to each row of the ANOVA table. Having prepared the ANOVA table, we are then able to determine whether the assignable sources of variation contribute a meaningful role in the underlying model for the design we are analyzing. The six ANOVA routines which are included in this package are probably the most widely used ANOVA procedures; because each routine offers analyses for several different forms of the data set, the six ANOVA routines really are able to analyze a wide range of designs.

This package uses the Basic Statistics and Data Manipulation package as the "front-end" routines to provide flexible data management prior to using the ANOVA routines. Moreover, the four auxiliary routines (orthogonal polynomials, treatment contrasts, multiple comparisons, and interaction plots) offer the necessary follow-up procedures from the ANOVA. Using one or more of these four routines and the new Response key, you can complete an ANOVA on one response and then easily carry out the same analysis on another response without having to specify the design structure again.

Since ANOVA is one of the most popular procedures used in the Statistical Laboratory we are sure that these routines will be used often on our campus. We hope you will find them as useful in your environment.

Programmers for this package are Conrad Schlundt, Chih-Ming Wang, and James Calvin, all of the Statistical Laboratory.

Thomas J. Boardman
Professor-In-Charge
Statistical Laboratory
Colorado State Univesity
Fort Collins, Colorado 80523

General Information for 9835A

System Configuration

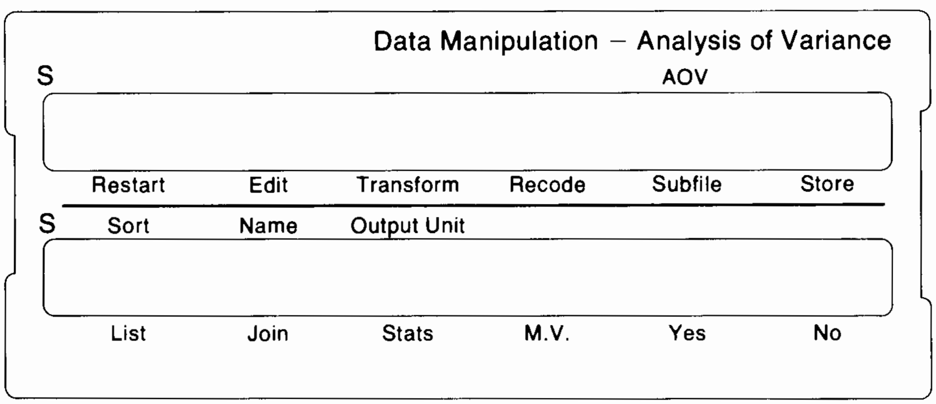
Necessary:

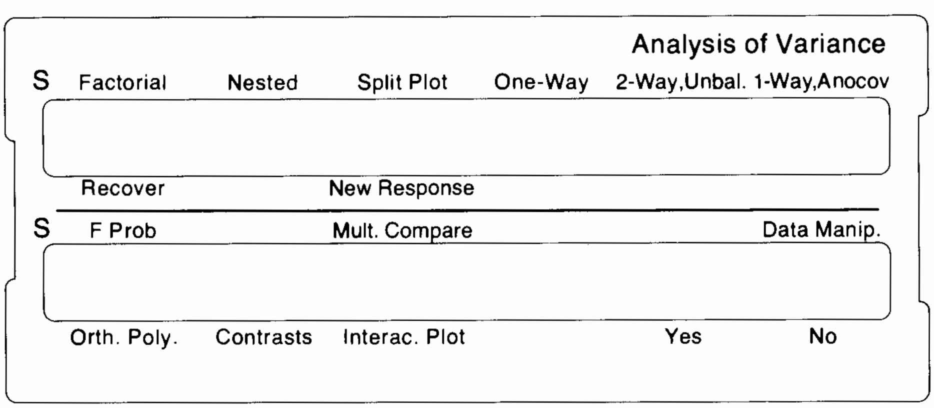
9835A Desktop Computer

Optional:

98337A – Plotter Graphics ROM
9872, 7245A, 7225A – External Plotters
9885M – Flexible Disk Drives (requires 98331A Mass Storage ROM)
2631A, 9876A, 9871A, 9866B, etc. – External Printers
9869A – Hopper Card Reader (requires 98332A General I/O ROM)

Special Function Keys





Part Numbers

The component parts of the Analysis of Variance package may be ordered separately:

Manual	09845-15171
Cartridges	09835-15074
	09835-15075

Key Overlays:

Data Manipulation – Analysis of Variance	7120-8883
Analysis of Variance	7120-8884

NOTE

1) For every place in the manual user instructions that "CONT" occurs, press "CONTINUE" instead.

2) Any references in this manual to the 9845B are also applicable to the 9835A.

Basic Stat and Data Manipulation

Introduction

Description

This set of programs allows you to enter a data matrix into memory and to then perform various operations on the data. The data entry may be made via keyboard, mass storage, card reader, or other devices. The operations on the data set include editing, transforming, recoding, sorting, creating subfiles, naming, storing and listing.

Features include a provision for missing data values, a provision for incorporating a subfile structure, the ability to store the data matrix and related information, the ability to join two data sets, error detection, the ability to correct many possible errors, and YES/NO keys to speed program use.

More specific objectives and features are listed with the instructions for the individual programs.

Special Considerations

Data Matrix Configuration

The data matrix incorporated in this program should be thought of as a p-by-n array whose columns correspond to observations and whose rows correspond to variables as shown below.

OBSERVATIONS

		O_1	O_2	O_3	...	O_n
	V_1				...	
	V_2				...	
VARIABLES	V_3				...	
	.	.	.	.		.
	.	.	.	.		.
	.	.	.	.		.
	V_p				...	

Subfiles may be created, in which case the structure becomes only slightly more complex as shown below.

		SUBFILE 1	SUBFILE 2		SUBFILE S
		$O_1\ O_2 \ldots O_{n_1}$	$O_{n_1+1}\ O_{n_1+2} \ldots O_{n_1+n_2}$	. . .	$O_{n_1+\ldots n_{s-1}+1} \ldots O_{n_1+\ldots+n_s}$
	V_1				
	V_2				
	.				
VARIABLES	.				
	.				
	V_p				

Limitations

The programs in this package have been designed to run on any 9845B regardless of memory size. They will operate on a maximum of 50 variables and a maximum of 20 subfiles. A total of 1500 data values may be input. So, for example, if the data set has two variables, 750 observations may be input.

Data File Configuration

The scratch file on the program medium, "DATA", and any files created to hold stored data and related information are configured as follows.

The data file is broken into logical records of 1280 bytes each. The first logical record is a "header file", which contains information pertinent to the data set stored in the remaining logical records. The header file contains the following information (variables):

- data set title (T$)
- number of observations (No)
- number of variables (Nv)
- variable names (Vn$(X))
- number of subfiles (Ns)
- subfile names (Sn$(*))
- subfile characterizations (Sc(*))

The remaining logical records contain D(*,*) — the data matrix.

Missing Values

If a data value is missing, it can be so designated by pressing the special function key labeled M.V., when entering the data from the keyboard. The number used to designate a missing value is −9999999.99999. The justification for this number is that (besides seeming unlikely to occur as a legitimate data point) it is easily picked out in a listing of the data. If the input is to be via cards, it may be more desirable to designate a missing value by a more easily typed number, for example, by 0 if zero is not a legitimate data point. The zeros could then be converted to the missing value recognized by the programs via a user-defined transformation.

Incorrect Responses

If a response outside the range of plausible responses is input from the keyboard, a message so stating is displayed for about three seconds. Program execution is resumed by asking the question or a previous question again.

If a plausible response is given, but it is not correct, one of three possibilities exists. First, if an incorrect value has been entered for a data point, it may be corrected using the EDIT program. Second, in many cases, responses to several questions are printed on the CRT and then a question such as "Is the above information correct?" is asked. This allows any of the printed information to be changed. Third, if a YES/NO question is answered incorrectly or if the above options are not offered, the program can be re-started by pressing STOP, then RUN.

Start

Object of Program

This program allows you to place a data matrix into memory. The data may be entered from the keyboard, from cards, or from some other input device. Conversely, the data may have been entered previously and stored in the program medium's scratch file ("DATA") or in a user created file on a tape cartridge, flexible disk or hard disk. In this case, the function of this program is to retrieve the previously stored data and place it into memory so that further operations can be performed. After the data is in memory, a listing option is provided.

Special Considerations

- The prompts concerning the data medium and program medium may cause confusion. The word "medium" is used since the set of programs making up the "BASIC STATISTICS AND DATA MANIPULATION" package may be tape, floppy disk or hard disk. Thus, the "program medium" refers to either the cartridge or the disk on which the programs making up this package are stored. Conversely, the "data medium" refers either to the cartridge or to the disk on which the file containing the data matrix resides. In some cases, the program medium and the data medium may be the same. However, this cannot be determined by the program and hence, the prompts are still displayed to make sure the correct medium is in the correct device.
- When entering data from the keyboard, an option to enter data one case at a time is offered. The following example will serve to explain this feature. Suppose an investigator has collected four observations on each of three variables. He has the following data matrix:

		Variable		
		1	2	3
Observation	1	10	2	5
	2	11	2	6
	3	9	3	7
	4	9	2	6

He elects to enter the data one case at a time. Then, when the prompt "Observation #, all variables (separated by commas)=?" is displayed, he enters 10, 2, 5 and presses CONT, etc. This allows for quick entry of the data.

- Because of the wide variety of formats that could be used when entering data via the card reader or "other" device, no attempt was made to program in the necessary statements. It will be necessary for you to provide the statements before using the program. The following procedure should be followed:

  For card reader input:

  1. Type: LOAD "START"
  2. Press: EXECUTE
  3. Type: EDITLINE Card
  4. Press: EXECUTE
  5. Type in and store the appropriate statements for card reader input.

  For "other" input:

  1. Type: LOAD "START"
  2. Press: EXECUTE
  3. Type: EDITLINE Other
  4. Press: EXECUTE
  5. Type in and store the appropriate statements for "other" input.

User Instructions

Getting Started

1. With the machine turned on, insert the "ANALYSIS OF VARIANCE, CARTRIDGE 1" cartridge into the tape drive.

2. Load the program into memory:

   a. Type: LOAD "AUTOST"

   b. Press: EXECUTE

3. Press: RUN

4. When "Is 'DATA MANIPULATION-ANALYSIS OF VARIANCE' overlay placed on keys?" is displayed:

   a. Press: YES when the overlay is in place.

General Information for 9845

System Configuration

Necessary:

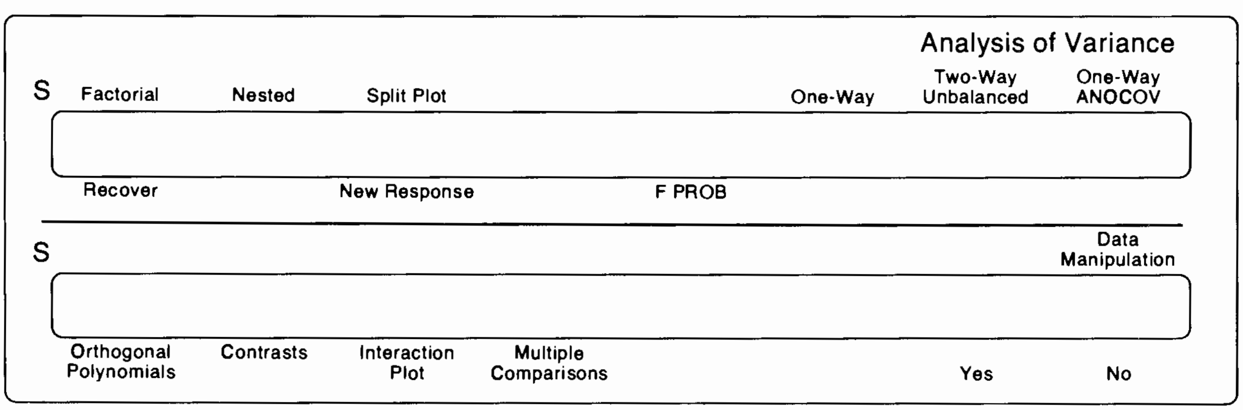
9845B or 9845C (Opt. 001) Desktop Computer
Option 560 – Thermal line printer (or an external printer)

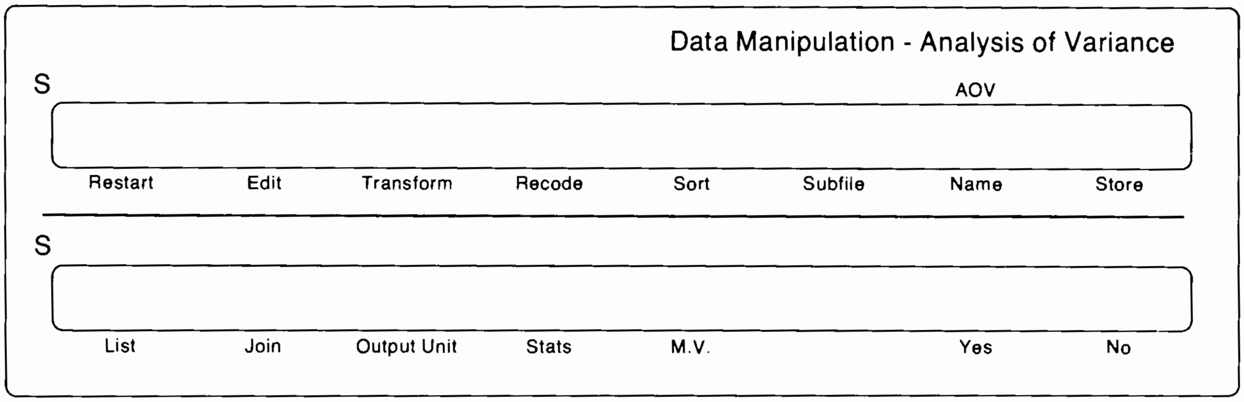
Optional:

Option 312	– I/O ROM for data entry from external devices
Option 311	– Graphics ROM (for 9845B only)
Option 600	– Secondary tape drive
Option 700	– Graphics Display Subsystem (for 9845B only)
9872	– External plotter
9885M	– Flexible Disk Drive (requires Option 313 – Mass Storage ROM)
7900 Series	– Hard Disc (requires Option 313 – Mass Storage ROM and 98041A Disc Interface)

Special Function Keys

Most of the operations are selected by pressing the appropriate special function key, shown below. Two of the keys those labeled Yes and No, may be used as responses to any question requiring a yes or no answer.





Part Numbers

The component parts of the ANALYSIS OF VARIANCE package may be ordered separately:

Manual		09845-15171
Cartridges		09845-15174
		09845-15175
Key Overlays:	Data Manipulation-Analysis of Variance	7120-7846
	Analysis of Variance	7120-7845

NOTE

Any references in this manual to 9845B are also applicable to the 9845C.

Hard-Copy Options

5. When "HARD-COPY OPTIONS" is printed and "Option number=?" is displayed:

   a. Enter: 1, if no hard copy is desired.

   b. Press: CONT

   c. Go to step 9.

   or

   *a. Enter: 2, if hard copy is desired from the machines's internal printer.

   b. Press: CONT

   c. Go to step 9.

   or

   a. Enter: 3, if hard copy is desired from an external printer.

   b. Press: CONT

6. When "Are you using an HP-IB Printer?" is displayed:

   a. Press: YES if an HP-IB Printer is being used.

   b. Go to step 7.

   or

   a. Press: NO if the printer is not HP-IB.

   b. Go to step 8.

7. When "Printer select code, bus address=?" [for example, 7,1] is displayed:

   a. Enter the select code, bus address for the HP-IB printer.

   b. Press: CONT

   c. Go to step 9.

8. When "Printer select code=?" is displayed:

   a. Enter the select code of the printer desired for hard copy.

   b. Press: CONT

*When using a 9835A, ignore option 2 a,b,c for an internal printer, and enter 2 for hard copy from an external printer.

Data Entry Modes

9. When "DATA ENTRY MODES" is printed and "Mode number=?" is displayed:

   a. Enter: 1, if data is to be entered from the keyboard.
   b. Press: CONT
   c. Go to step 10.

   or

   a. Enter: 2,if data is to be entered from a cartridge or a disk.
   b. Press: CONT
   c. Go to step 15.

   or

   a. Enter: 3, if data is to be entered via cards.
   b. Press: CONT
   c. Go to step 10.

   or

   a. Enter: 4, if data is to be entered from some other device.
   b. Press: CONT

Creating the Data Set

10. When "Project title for this data set (<= 80 characters)=?" is displayed:

    a. Enter up to 80 characters to be used as the name of the data set.
    b. Press: CONT

11. When "Number of variables=?" is displayed:

    a. Enter the number of variables you want in the data set.
    b. Press: CONT

12. When "Number of observations/variable=?" is displayed:

    a. Enter the number of observations for each variable in the data set.
    b. Press: CONT

13. When "VARIABLE NAMES" is printed and "Variable #I name (<= 10 characters)=?" is displayed:

    a. Enter the name of the Ith variable - it may be up to 10 characters long.
    b. Press: CONT
    c. Repeat steps a and b until all variables have been named, then go to step 14.

14. When "Is above information correct" is displayed:

    a. Press: YES if the variable names are as desired.

    b. Go to step 21.

    or

    a. Press: NO if changes to the variable names are desired.

    b. Go to step 13.

Using an Existing Data Set

15. When "Is data stored on the program medium's scratch file (DATA)?" is displayed:

    a. Press: YES if the data is stored in "DATA", the scratch file on the program tape.

    b. The data and related information are loaded into memory at this point.

    c. Go to step 21.

    or

    a. Press: NO if the data is stored on a user-created file.

16. When file name specifications are printed and "Data file name=?" is displayed:

    a. Enter the data file name (less than or equal to six characters) followed by a colon and the mass storage unit specifier.

    b. Press: CONT

17. When "Was data stored by this program?" is displayed:

    a. Press: YES if the data file was created by the "BASIC STATISTICS AND DATA MANIPULATION" routine or by any routine which stores a header file with the data (see Special Considerations).

    b. Go to step 18.

    or

    a. Press: NO if the file was created by a program which does not store header information along with the data.

    b. Go to step 10.

18. When "Is data medium placed in device?" is displayed:

    a. Press: YES when the data medium is in place.

19. The data and related information are loaded into memory at this point.

20. When "Is program medium replaced in device?" is displayed:

    a. Press: YES when the program medium is in place (or if it was never removed).

Summary of the Data Set

21. At this point a summary of the data set is printed.
22. If the input mode is via keyboard, card reader or "OTHER", go to step 25.
23. When "List data?" is displayed:

    a. Press: YES if a hard-copy listing of the data is desired.

    b. Go to step 1 of the user instructions for the LIST program.

    or

    a. Press: NO if a hard-copy listing is not desired.

Selecting a Program

24. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

    b. Go to the user instructions for the selected key.

Data Input

25. At this point the subprogram "Input" is called. A subprogram is used here for the sole purpose of variable dimensioning.
26. When "DATA INPUT" is printed:

    a. If the data is being entered from the keyboard, go to step 32.

    or

    a. If the data is being entered from cartridge or disk, go to step 32.

    or

    a. If the data is being entered via cards, go to step 40.

    or

    a. If the data is being entered from some other device, go to step 41.

27. When the program notes are printed:

    a. If there is only one variable in the data set, go to step 28.

    or

    a. If there are two or more variables in the data set, go to step 29.

28. When the heading is printed and "Observation #I" is displayed:

    a. Enter the value of the Ith observation.

b. Press: CONT

c. Repeat steps a and b until all observations have been entered.

d. Go to step 42.

29. When "Enter data one case at a time, that is, by observation?" is displayed:

 a. Press: YES if you desire to key in the values of all variables for a given observation at once.

 b. Go to step 31.

 or

 a. Press: NO if you desire to key in the value of one variable at a time for a given observation.

30. When "Observation #I, Variable #J" is displayed:

 a. Enter the value of the Ith observation, Jth variable.

 b. Press: CONT

 c. Repeat steps a and b until all variables for all observations are entered.

 d. Go to step 42.

31. When "Observation #I, all variables (separated by commas)=?" is displayed:

 a. Enter variable 1, comma, variable 2, comma, . . ., variable P for Observation #I.

 b. Press: CONT

 c. Repeat steps a and b until all variables for all observations are entered.

 d. Go to step 42.

32. When "Are MISSING VALUES denoted by −9999999.99999?" is displayed:

 a. YES if missing values were specified by pressing MISSING VALUE (M.V.) when the data was originally entered, or if there are no missing values.

 b. Go to step 34.

 or

 a. Press: NO if missing values are specified by some other value.

33. When "Missing value=?" is displayed:

 a. Enter the value that has been used to specify a missing data point.

 b. Press: CONT

34. When "Is data in proper configuration, that is, variables = rows, observations = columns?" is displayed:

    a. Press: YES if the data matrix contains a variable in each row and an observation in each column.

    b. Go to step 36.

    or

    a. Press: NO if the data matrix is in a different configuration.

35. When "Data stored as contiguous array with observation = rows, variables = columns?" is displayed:

    a. Press: YES if the data matrix contains an observation in each row and a variable in each column.

    b. Go to step 36.

    or

    a. Press: NO if the data is configured in a different manner.

    b. Execution stops - the data set is unacceptable for use by this program and therefore must be keyed in.

36. When "Is data medium placed in device?" is displayed:

    a. Press: YES when data medium is in place.

37. At this point the data is loaded into memory.

38. When "Is program medium replaced in device?" is displayed:

    a. Press: YES when program medium is in place (or if it was never removed).

39. At this point the data is stored on the scratch file of the program medium (DATA). Go to step 42.

40. See Special Considerations.

41. See Special Considerations.

42. When "List data?" is displayed:

    a. Press: YES if a hard-copy listing of the data is desired.

    b. Go to step 1 of the user instructions for LIST.

    or

    a. Press: NO if no listing is desired.

Selecting a Program

43. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys defined on the "BASIC STATISTICS AND DATA MANIPULATION" overlay.

    b. Go to step 1 of the user instructions for the selected key.

Example

The data listed below provides an example of most of the operations that can be performed by this package. There are five variables:

1. average monthly temperature (°C)
2. monthly production (kg.)
3. number of working days
4. number of people on the payroll
5. monthly water use (litres)

There are 17 observations, one per month.

```
*****************************************************************************************
*                                  DATA MANIPULATION                                    *
*****************************************************************************************
```

SAMPLE

Data file name: DATA
Number of observations: 17
Number of variables: 5

Variables names:
1. Temp(C)
2. Production
3. Days
4. Payroll
5. Water Use

Subfiles: NONE

SAMPLE

OBS#	Variable # 1	Variable # 2	Variable # 3	Variable # 4	Variable # 5
1	14.90000	6396.00000	21.00000	134.00000	3373.00000
2	18.40000	5736.00000	22.00000	146.00000	3110.00000
3	21.60000	6116.00000	22.00000	158.00000	3180.00000
4	25.20000	8287.00000	20.00000	171.00000	3293.00000
5	26.30000	13313.00000	25.00000	198.00000	3390.00000
6	27.20000	13108.00000	23.00000	194.00000	4287.00000
7	22.20000	10768.00000	20.00000	180.00000	3852.00000
8	17.10000	12173.00000	23.00000	191.00000	3366.00000
9	12.50000	11390.00000	20.00000	195.00000	3532.00000
10	6.90000	12707.00000	20.00000	192.00000	3614.00000
11	6.40000	15022.00000	22.00000	200.00000	3896.00000
12	13.30000	13114.00000	19.00000	211.00000	3437.00000
13	18.20000	12257.00000	22.00000	203.00000	3324.00000
14	22.80000	13118.00000	22.00000	197.00000	3214.00000
15	26.10000	13100.00000	21.00000	196.00000	4345.00000
16	26.30000	16716.00000	21.00000	205.00000	4936.00000
17	4.20000	14056.00000	22.00000	205.00000	3624.00000

Edit

Object of Program

This program is designed to allow you to perform a variety of editing procedures on the data matrix. The editing capabilities include: the changing of an incorrect data value, the deletion of a variable, the deletion of an observation, the addition of an observation, the insertion of an observation (if the data is ordered), and the addition of a variable. All of the above operations can be performed repeatedly; for example, three variables could be deleted in succession. After the data matrix has been edited, you are given the option of listing the edited data.

Special Considerations

Order of Corrections

As stated in the program note printed on the screen, the data is renumbered after deletions or insertions are performed. For this reason, if more than one deletion (insertion) is to be performed, it is recommended that the highest-numbered observation (or variable) be deleted, then the next highest-numbered, etc. For example, if observations three and eight are to be deleted, then it is recommended to delete observation eight first, then observation three. Notice that if observation three were deleted first, the subsequent renumbering would move observation eight to position seven. The recommendation is meant to alleviate confusion which may occur due to the renumbering.

Subfiles

Insertions or deletions of observations will affect the content of subfiles which are in existence at the time of editing; for example, if subfile one consists of the first 10 observations while subfile two consists of the last 20 and if observation five is deleted, then observation ten (formerly numbered 11) will have jumped from subfile two to subfile one. Thus, it may be necessary to change the subfile structure after editing. It is recommended that subfiles be created only after all editing has been performed.

Methods and Formulae

The data matrix is redimensioned into a row vector to facilitate the shuffling of elements necessitated by the editing operations. The vector contains all the observations of variable one, followed by the observations of variable two, etc. When an observation is inserted, for example, the elements of the data vector are shuffled one at a time to make room for the incoming observation. Similarly, when an observation is deleted, the remaining observations are "packed" together so that the resultant data vector has no "holes" between observations.

User Instructions

Press: The special function key labeled Edit.

Correcting a Data Value

1. When "Correct a data value?" is displayed:

   a. Press: NO if there are no data values present that need to be changed.

   b. Go to step 6.

   or

   a. Press: YES if there is a data point in memory that needs to be changed.

2. When "Observation number=?" is displayed:

   a. Enter the observation number of the data value that is to be changed.

   b. Press: CONT

   c. If there is only one variable in the data set, go to step 4.

3. When "Variable number=?" is displayed:

   a. Enter the variable number of the data value that is to be changed.

   b. Press: CONT

4. When "Old value='.' −correct value=?" is displayed:
   (The old value is displayed so one can be sure that the correction is being made as anticipated.)

   a. Enter the correct data value.

   b. Press: CONT

   c. A hard-copy note of the change is printed.

5. When "Correct another value?" is displayed:

   a. Press: YES if there is another data point in memory that needs to be changed.

   b. Go to step 2.

   or

   a. Press: NO if there are no more changes to be made to existing data points.

6. If there is only one variable in the data set, go to step 11.

Deleting a Variable

7. When "Delete a variable?" is displayed:

   a. Press: NO if no variables are to be removed from the data set.

   b. Go to step 11.

   or

   a. Press: YES if you desire to remove one or more variables from the data set.

8. When the program note is printed, and "Number of the variable to be deleted=?" is displayed:

   a. Enter the highest-numbered variable to be deleted. For example, if variables 2 and 7 are to deleted, enter 7 first (since the variables are renumbered after a deletion, and if 2 were deleted first, then variable 7 would be referred to as variable 6).

   b. Press: CONT

   c. A hard-copy note of the deletion is printed.

9. If there is only one variable remaining in the data set, go to step 11.

10. When "Delete another variable?" is displayed:

    a. Press: YES if you desire to remove another variable from the data set.

    b. Go to step 8.

    or

    a. Press: NO if no more variables are to be deleted from the data set.

Deleting an Observation

11. When "Delete an observation?" is displayed:

    a. Press: NO if no observations are to be removed from the data set.

    b. Go to step 14.

    or

    a. Press: YES if you desire to remove one or more observations from the data set.

12. When the program note is printed and "Number of the observation to be deleted=?" is displayed:

    a. Enter the highest-numbered observation to be deleted; for example, if observations 2 and 5 are to be deleted, enter 5 first (since the observations are renumbered after each deletion, if 2 were deleted first, then observation 5 would become observation 4).

    b. Press: CONT

    c. A hard-copy note of the deletion is printed.

13. When "Delete another observation?" is displayed:

    a. Press: YES if it is desired to remove another observation from the data set.

    b. Go to step 12.

    or

    a. Press: NO if no more observations are to be removed from the data set.

Adding an Observation

14. If the addition of one more observation will exceed available memory, go to step 23.

15. When "Add an observation?" is displayed:

    a. Press: NO if no observations are to be added to the end of the data set or inserted into the data set.

    b. Go to step 23.

    or

    a. Press: YES if you desire to add more observations to the end of the data set or to insert observations into the data set.

16. When "Are observations ordered, that is, should additions be inserted?" is displayed:

    a. Press: YES if the data follows some type of order and if you desire to insert an observation, for example, between observations 4 and 5.

    b. Go to step 19.

    or

    a. Press: NO if the observations are to be added to the end of the data set.

17. When "How many observations are to be added?" is displayed:

    a. Enter the number of observations to be added to the "end" of the data set.

    b. Press: CONT

18. When "Observation #'I', Variable #'J'=?" is displayed:

    a. Enter the data point corresponding to the Jth variable of the Ith observation.

    b. Press: CONT

    c. A hard-copy note of the addition is printed.

    d. Repeat steps a-c until data has been entered for all variables of each of the observations added.

    e. Go to step 23.

19. When the program note is printed and "Insertion to precede observation #?" is displayed:

    a. Enter the number of the observation which the insertion will precede. For example, if an observation is to be inserted between observations 8 and 9, enter 9. Since the observations are renumbered after each insertion, insertions should be made between the highest-numbered observations first. For example, if insertions are to be made between observations 4 and 5 as well as between observation 8 and 9, the latter should be performed first, that is, 9 should be entered first.

    b. Press: CONT

20. When "Observation #'I', Variable #'J'=?" is displayed:

    a. Enter the data point corresponding to variable J of the inserted observation.

    b. Press: CONT

    c. A hard-copy note of the insertion is printed.

    d. Repeat steps a-c until data has been entered for all variables.

21. If the addition of one more observation will exceed program limitations, go to step 23.

22. When "Insert another observation?" is displayed:

    a. Press: YES if you desire to insert another observation.

    b. Go to step 19.

    or

    a. Press: NO if no more observations are to be inserted.

Adding a Variable

23. If the addition of a variable will exceed program limitations, go to step 29.

24. When "Add a variable?" is displayed:

    a. Press: NO if no variables are to be added from the keyboard.

    b. Go to step 29.

    or

    a. Press: YES if you desire to add one or more variables from the keyboard.

25. When "Variable name (<= 10 characters)" is displayed:

    a. Enter the name of the variable to be added - it may be up to 10 characters long.

    b. Press: CONT

26. When "Variable #'I', Observation #'J'=?" is displayed:

    a. Enter the Jth observation of the variable being added.
    b. Press: CONT
    c. A hard-copy note of the addition is printed.
    d. Repeat steps a-c until all observations have been entered for the added variable.

27. If the addition of one more variable will exceed available memory, go to step 29.
28. When "Add another variable?" is displayed:

    a. Press: YES if you desire to add another variable from the keyboard.
    b. Go to step 25.

    or

    a. Press: NO if no more variables are to be added from the keyboard.

29. When "More corrections?" is displayed:

    a. Press: YES if more corrections to the data matrix are required.
    b. Go to step 1.

    or

    a. Press: NO if no more corrections to the data set are necessary.

30. At this point the corrected data matrix and related information are recorded in file "DATA" of the program medium.
31. "List data?" is displayed:

    a. Press: YES if a hard-copy listing of the corrected data set is desired.
    b. Go to step 1 of the user instructions for LIST.

    or

    a. Press: NO if a hard-copy listing of the data is not required.

32. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.
    b. Go to step 1 of the user instructions for the selected key.

Example

The program EDIT was run in the print all mode. The printout shows a correction, deletion of an observation, and addition of an observation. The edited data was then listed.

```
Correct a data value?
"YES"
Observation number = ?
11
Variable number = ?
2
Old value =  15022 -- Correct value =
?
15024
Observation # 11 Variable # 2 -- correct value =  15024
Correct another value?
"NO"

Delete a variable?
"NO"
Delete an observation?
"YES"
Number of the observation to be deleted = ?
10
Observation # 10 has been deleted, 16 observations remain.
Delete another observation?
"NO"

Add an observation?
"YES"
Are observations ordered, i.e., should additions be inserted?
"NO"
How many observations are to be added?
1
Observation # 17 Variable # 1  =
?
4.2
Observation # 17 Variable # 1  =  4.2
Observation # 17 Variable # 2  =
?
12707
Observation # 17 Variable # 2  =  12707
Observation # 17 Variable # 3  =
?
20
Observation # 17 Variable # 3  =  20
Observation # 17 Variable # 4  =
?
192
Observation # 17 Variable # 4  =  192
Observation # 17 Variable # 5  =
?
3614
Observation # 17 Variable # 5  =  3614

  Total number of observations now =  17

Add a variable?
"NO"
More corrections?
"NO"
List data?
"YES"
```

SAMPLE

OBS#	Variable # 1	Variable # 2	Variable # 3	Variable # 4	Variable # 5
1	14.90000	6396.00000	21.00000	134.00000	3373.00000
2	18.40000	5736.00000	22.00000	146.00000	3110.00000
3	21.60000	6116.00000	22.00000	158.00000	3180.00000
4	25.20000	8287.00000	20.00000	171.00000	3293.00000
5	26.30000	13313.00000	25.00000	198.00000	3390.00000
6	27.20000	13108.00000	23.00000	194.00000	4287.00000
7	22.20000	10768.00000	20.00000	180.00000	3852.00000
8	17.70000	12173.00000	23.00000	191.00000	3366.00000
9	12.50000	11390.00000	20.00000	195.00000	3532.00000
10	6.40000	15024.00000	22.00000	200.00000	3896.00000
11	13.30000	13114.00000	19.00000	211.00000	3437.00000
12	18.20000	12257.00000	22.00000	203.00000	3324.00000
13	22.80000	13118.00000	22.00000	197.00000	3214.00000
14	26.10000	13100.00000	21.00000	196.00000	4345.00000
15	26.30000	16716.00000	21.00000	205.00000	4936.00000
16	4.20000	14056.00000	22.00000	205.00000	3624.00000
17	4.20000	12707.00000	20.00000	192.00000	3614.00000

Transform

Object of Program

This program allows you to transform one or two variables in the data matrix via 16 pre-specified functions or through a function which you specify. The transformed data may then be treated as a new variable, or it may replace the elements of an existing variable. Hence, transformations on more than two variables may be made iteratively or via a user-defined transformation.

The transformations available are:

1. $a X^{b}+c$
2. $a \log(bX)+c$
3. $a \ln(bX)+c$
4. $a \exp(bX)+c$
5. $a (b^{cX})$
6. $a \cos(bX)+c$
7. $a \sin(bX)+c$
8. $a+\sqrt{\arcsin(bX)+c}$
9. $aX+bY+c$
10. $aX^{b}Y^{c}$
11. $a \log(bX+cY)$
12. $a \ln(bX+cY)$
13. $a \cos(bX+cY)$
14. $a \sin(bX+cY)$
15. PROUND(X,a) [round to specified power of 10]
16. DROUND(X,a) [round to specified no. of digits]
17. User Defined

Special Considerations

Missing Values

None of the 16 pre-specified transformations are applied to missing values. Thus, missing values are unaffected by these transformations. However, this is not necessarily the case with the user-defined transformation. If you define a transformation and there are missing values, you must make provisions to ensure that the transformation is not applied to the missing values (unless, of course, this is desired). This may be accomplished as explained below.

User-Defined Transformation

Up to 10 lines of code may be used to define a transformation, namely lines 1330 through 1339. These lines may be typed in and stored successively prior to pressing 'CONT'. (See user instructions.) The following example shows the form of a typical user-defined transformation. Suppose the data set consists of four variables with 20 observations each. There are missing values, which are to be left unchanged. You desire to form variable five as the sum of the exponentials of variables one and three. The following sequence should be carried out at step three of the TRANSFORM User Instructions.

a. Type: 1330 D(Z,I)=EXP(D(1,I))+EXP(D(3,I))

b. Press: STORE

c. Type: 1331 IF D(1,I)=-9999999.99999 OR D(3,I)=-9999999.99999 THEN D(5,I)=1

d. Press: STORE

e. Press: CONT

The variable Z is used to identify variable in which the result of the transformation is to be stored. Notice that the elements of D(.,.) are the data - the first subscript refers to variable number while the second subscript refers to observation number (and should always be 'I' here). Line 1331 sets D(5,I)=1 if a missing value is present in either of variables one or three. Notice that D(Z,I) or D(5,I) may be used interchangeably since Z has been assigned the value of 5 previously in the program.

User Instructions

Press: The special function key labeled Transf

1. When the program notes and transformations are printed and "Transformation number=?" is displayed:

    a. Enter the number of the transformation which is desired (1-17, inclusive).

    b. Press: CONT

2. If the selected transformation was not 'User defined', that is, not transformation number 17, go to step 4.

3. When the program notes are printed and "Ready to continue?" is displayed:

    a. Type: "1330 D(Z,I)='defined transformation' "; the form of the defined transformation is discussed in Special Considerations.

    b. Press: STORE

    c. Press: YES when ready to continue.

    d. Go to step 11.

4. When the selected transformation is printed and "Number of the variable corresponding to X=?" is displayed:

a. Enter the variable number corresponding to X, where X refers to the variable in the selected transformation.

b. Press: CONT

5. If the selected transformation number is less than 9 or greater than 14, go to step 7.

6. When "Number of the variable corresponding to Y=?" is displayed:

a. Enter the variable number corresponding to Y, where Y refers to the variable in the selected transformation.

b. Press: CONT

7. When "Parameter a=?" is displayed:

a. Enter the value of parameter a, where a refers to the parameter in the selected transformation.

b. Press: CONT

8. If the selected transformation number is greater than 14, go to step 11.

9. When "Parameter b=?" is displayed:

a. Enter the value of parameter b, where b refers to the parameter in the selected transformation.

b. Press: CONT

10. When "Parameter c=?" is displayed:

a. Enter the value of parameter c, where c refers to the parameter in the selected transformation.

b. Press: CONT

11. When "Store transformed data in variable # (<='I')" is displayed:

a. Enter the number of the variable in which the transformed data is to be stored - it may be less than or equal to I.

b. Press: CONT

12. If the variable specified in step 11 existed previously, that is, if the transformed data is being stored in place of previously existing data, go to step 14.

13. When "Variable name (<= 10 characters)=?" is displayed:

a. Enter the name of the variable to contain the transformed data - it may be up to 10 characters long.

b. Press: CONT

14. If the transformation was 'User defined', go to step 16.

15. When "Is above information correct?" is displayed:

    a. Press: NO if a mistake has been made during entry of the information printed.

    b. Go to step 1.

    or

    a. Press: YES if no corrections are to be made.

16. At this point a hard-copy summary of the transformation is made (unless it was user defined), then the transformation is carried out.

17. When "More transformations?" is displayed:

    a. Press: YES if more transformations are desired.

    b. Go to step 1.

    or

    a. Press: NO if no more transformations are desired.

18. At this point the altered data matrix and related information are stored on file "DATA" of the program medium.

19. When "List data?" is displayed:

    a. Press: YES if a hard-copy listing of the data is desired.

    b. Go to step 1 of the user instructions for LIST.

    or

    a. Press: NO if a hard-copy listing of the data is not required.

20. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

    b. Go to step 1 of the user instructions for the selected key.

Example

Variable five, Water Use, was tranformed, converting litres to U.S. gallons. Transformation 1 was used, with a=0.2642, b=1,c=0.


```
Transformation number = ?
1
Number of the Variable number corresponding to X = ?
5
Parameter a = ?
.2642
Parameter b = ?
1
Parameter c = ?
0
Store transformed data in Variable # ( <= 6 )
?
5
Is above information correct?
"YES"
The following transformation was performed: a*(X^b)+c
  where a =  .2642
        b =  1
        c =  0
        X is Variable # 5
        Transformed data is stored in Variable # 5 (Water Use).


More transformations?
"NO"
List data?
"YES"
```

SAMPLE

OBS#	Variable # 1	Variable # 2	Variable # 3	Variable # 4	Variable # 5
1	14.90000	6396.00000	21.00000	134.00000	891.14660
2	18.40000	5736.00000	22.00000	146.00000	821.66200
3	21.60000	6116.00000	22.00000	158.00000	840.15600
4	25.20000	8287.00000	20.00000	171.00000	870.01060
5	26.30000	13313.00000	25.00000	198.00000	895.63800
6	27.20000	13108.00000	23.00000	194.00000	1132.62540
7	22.20000	10768.00000	20.00000	180.00000	1017.69840
8	17.70000	12173.00000	23.00000	191.00000	889.29720
9	12.50000	11390.00000	20.00000	195.00000	933.15440
10	6.40000	15024.00000	22.00000	200.00000	1029.32320
11	13.30000	13114.00000	19.00000	211.00000	908.05540
12	18.20000	12257.00000	22.00000	203.00000	878.20080
13	22.80000	13118.00000	22.00000	197.00000	849.13880
14	26.10000	13100.00000	21.00000	196.00000	1147.94900
15	26.30000	16716.00000	21.00000	205.00000	1304.09120
16	4.20000	14056.00000	22.00000	205.00000	957.46080
17	4.20000	12707.00000	20.00000	192.00000	954.81880

Recode

Object of Program

This program allows you to assign codes to various categories or classes of data. The categories are intervals along the real number line and 20 of these may be specified. The recoding is done on one variable at a time. The same coding scheme may be used iteratively on successive variables. A summary of the coding intervals, codes, and number of observations assigned to each code is printed as hard copy.

Special Considerations

Coding Schemes

Four coding schemes are available for the sole purpose of eliminating unnecessary entries from the keyboard. If the coding intervals are all of the same length and are contiguous, that is, together they form a connected interval, then the interval construction can be accomplished internally knowing only the interval length and lower limit for the first interval. Similarly, if the intervals are of equal length but noncontiguous, for example, [10,20), [25,35), [35,45), [50,60), then the lower limit of each interval needs to be specified but the upper limit may be computed internally. Hence, the coding schemes are meant only to minimize the amount of information which needs to be entered from the keyboard. Clearly, the coding intervals could all be constructed by requiring you to enter the lower and upper limits for each and every interval (which is necessary, and what is done if the intervals are unequal and non-contiguous).

Same Coding Scheme

The coding is carried out on one variable at a time. However, if you desire to code both variables one and two according to the same coding intervals, these intervals need to be contructed only once. A positive response to the option offered by "Use same coding scheme?" allows variable two to be coded according to the same scheme without constructing the intervals a second time. If, however, you desire to code variable two according to a different scheme, it is possible to construct a second set of coding intervals by giving a negative response to the above prompt.

Brackets

The brackets used to denote the coding intervals are meant to follow their usual mathematical interpretation, that is, the intervals are closed on the left and open on the right.

Observation not in an Interval

If an observation does not fall into any of the coding intervals, its value is not changed during the coding process.

User Instructions

Press: The special function key labeled Recode.

1. When the title is printed and "Store recoded data in variable # (<='I')?" is displayed:
   a. Enter the number of the variable in which the recoded data is to be stored - it may be less than or equal to I.
   b. Press: CONT
2. If the variable specified in step 1 existed previously, that is, if the recorded data is being stored in place of previously existing data, go to step 4.
3. When "Variable name (<= 10 characters)=?" is displayed:
   a. Enter the name of the variable which will contain the recoded data - it may be up to 10 characters long.
   b. Press: CONT
4. When "Number of the variable to be recoded=?" is displayed:
   a. Enter the number of the variable which is to be recoded.
   b. Press: CONT
5. When the data coding schemes are printed and "Option number=?" is displayed:
   a. Enter the option number associated with the desired recoding scheme.
   b. Press: CONT
6. When "Number of recoding intervals to be specified (<=20)=?" is displayed:
   a. Enter the number of categories into which the data will be divided, or equivalently, the number of codes which will be assigned.
   b. Press: CONT
7. If the specified coding scheme does not have equal intervals, go to step 9.
8. When "Length of each interval=?" is displayed:
   a. Enter the common increment to be used to generate the intervals.
   b. Press: CONT
9. If the data recoding scheme is non-contiguous, go to step 11.
10. When "Lower limit of first interval=?" is displayed:
   a. Enter the smallest number which will be assigned code #1.
   b. Press: CONT

11. Note: The following steps, 12 through 14, are repeated the number of times specified in step 6 until all information is input for each of the recoding intervals. Some of the steps are skipped for certain schemes, since the necessary information is computed internally. For scheme #1, only step 14 needs to be repeated. For scheme #2, only steps 13 and 14 need to be repeated. For scheme #3, only steps 12 and 14 need to be repeated. For scheme #4, steps 12-14 need to be repeated.

12. When "Lower limit of interval #'I'=?" is displayed:

    a. Enter the smallest number which will be assigned the code associated with interval #I.

    b. Press: CONT

13. When "Upper limit of interval #'I'=?" is displayed:

    a. Enter the upper bound of numbers which will be assigned the code associated with interval #I.

    b. Press: CONT

14. When "For data falling in interval #'I', code=?" is displayed:

    a. Enter the code which will be assigned to those data values falling in the Ith coding interval.

    b. Press: CONT

    c. Repeat as noted in instruction 11.

15. When "Is above information correct?" is displayed:

    a. Press: NO if a mistake has been made in entering the information.

    b. Go to step 5.

    or

    a. Press: YES if the recoding scheme is as desired.

16. At this point the recoding is carried out and a hard-copy note of the coding is printed.

17. When "Recode more data?" is displayed:

    a. Press: YES if you desire to recode more data.

    b. Go to step 18.

    or

    a. Press: NO if no more recoding is desired.

    b. Go to step 19.

18. When "Use same coding scheme?" is displayed:

    a. Press: YES if the same intervals and codes are desired, that is, if the identical recoding scheme is to be applied to another variable.

    b. Go to steps 1-4, then directly to 16.

    or

    a. Press: NO if a different recoding scheme is desired.

    b. Go to step 1.

19. At this point the altered data matrix and related information are stored on file "DATA" of the program medium.

20. When "List data?" is displayed:

    a. Press: YES if a hard-copy listing of the data is desired.

    b. Go to step 1 of the user instructions for LIST.

    or

    a. Press: NO if no listing is desired.

21. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

    b. Go to step 1 of the user instructions for the selected key.

Example

Using the original data, variable 5, Water Use, was recoded, and the resulting data was stored as a new variable, 6. The purpose of the recoding operation was to simplify the data, replacing entire intervals by single numbers.

```
Store recoded data in Variable # (<= 6 )
?
6
Variable name (<= 10 characters) = ?
Coded Use
Number of the variable to be recoded = ?
5
Option number = ?
1
Number of recoding intervals to be specified (<=20) = ?
5
Length of each interval = ?
400
Lower limit of first interval = ?
3000
For data falling in interval 1 , code =
?
30
For data falling in interval 2 , code =
?
34
For data falling in interval 3 , code =
?
38
For data falling in interval 4 , code =
?
42
For data falling in interval 5 , code =
?
46
Is above information correct?
"YES"
```

```
Variable # 5 is recoded into 5 categories, and the recoded
values are stored in Variable # 6 , where:

              CATEGORY BOUNDS              # OBS
           LOWER             UPPER         CODED         CODE
        3000.000          3400.000             8       30.000
        3400.000          3800.000             4       34.000
        3800.000          4200.000             2       38.000
        4200.000          4600.000             2       42.000
        4600.000          5000.000             1       46.000

Recode more data?
"YES"
Use same coding scheme?


Recode more data?
YES
```

```
Recode more data?
NO
```

SAMPLE

OBS#	Variable # 1	Variable # 2	Variable # 3	Variable # 4	Variable # 5	Variable # 6
1	14.90000	6396.00000	21.00000	134.00000	3373.00000	30.00000
2	18.40000	5736.00000	22.00000	146.00000	3110.00000	30.00000
3	21.60000	6116.00000	22.00000	158.00000	3180.00000	30.00000
4	25.20000	8287.00000	20.00000	171.00000	3293.00000	30.00000
5	26.30000	13313.00000	25.00000	198.00000	3390.00000	30.00000
6	27.20000	13108.00000	23.00000	194.00000	4287.00000	42.00000
7	22.20000	10768.00000	20.00000	180.00000	3852.00000	38.00000

8					
	17.70000	12173.00000	23.00000	191.00000	3366.00000
	30.00000				
9					
	12.50000	11390.00000	20.00000	195.00000	3532.00000
	34.00000				
10					
	6.90000	12707.00000	20.00000	192.00000	3614.00000
	34.00000				
11					
	6.40000	15022.00000	22.00000	200.00000	3896.00000
	38.00000				
12					
	13.30000	13114.00000	19.00000	211.00000	3437.00000
	34.00000				
13					
	18.20000	12257.00000	22.00000	203.00000	3324.00000
	30.00000				
14					
	22.80000	13118.00000	22.00000	197.00000	3214.00000
	30.00000				
15					
	26.10000	13100.00000	21.00000	196.00000	4345.00000
	42.00000				
16					
	26.30000	16716.00000	21.00000	205.00000	4936.00000
	46.00000				
17					
	4.20000	14056.00000	22.00000	205.00000	3624.00000
	34.00000				

Sort

Object of Program

This program allows the data matrix, or subfiles thereof, to be sorted according to the values of one variable. For example, suppose an investigator has five observations of three variables, say height, weight and age and wanted to arrange the observations in ascending order according to age. This is accomplished by sorting the data matrix according to variable three.

Special Considerations

Subfile Structure Options

If subfiles are ignored, the entire data set will be sorted and, in the process, the composition of the subfiles is subject to change. The option of sorting certain subfiles may be used to sort a single subfile or a set of successive subfiles according to one variable. The option of sorting all subfiles may be used to sort every subfile. The options of sorting certain subfiles and sorting all subfiles treat each subfile as if it were a separate data set. Thus, the sort is done with respect to one subfile at a time.

What Happens

It is important to note that entire observations are moved when the sort is carried out. Thus, referring to the example given in Object of Program, a person's height and weight remain with the person's age as shown below.

Original Matrix

		Height	Weight	Age
	1	72	170	21
	2	70	165	25
OBSERVATION	3	69	150	20
	4	70	165	25
	5	73	160	19

Matrix Sorted By Age

VARIABLE

		Height	Weight	Age
	1	73	160	19
	2	69	150	20
OBSERVATION	3	72	170	21
	4	70	165	25
	5	70	165	25

User Instructions

Press: The special function key labeled Sort.

1. At this point the subprogram "Sort" is called. A subprogram is used here for the sole purpose of variable dimensioning.
2. When "Number of the Variable on which to sort=?" is displayed:

   a. Enter the number of the variable on which the matrix will be sorted; that is, the variable whose observations will be arranged in ascending order.

   b. Press: CONT

3. If the data set contains no subfiles, go to step 7.
4. When the subfile structure options are printed and "Option number=?" is displayed:

   a. Enter 1 if you desire to ignore the subfile structure, that is, to lump all the observations together for the sort. Note that this may shift observations from subfile to subfile.

   b. Press: CONT

   c. Go to step 7.

   or

   a. Enter 2 if you desire to sort a set of one or more successive subfiles, for example, to sort subfiles 3 through 5. The sort will be done by subfile: subfile 3 will be sorted, then 4, and finally 5.

   b. Press: CONT

   c. Go to step 5.

   or

   a. Enter 3 if you desire to sort all of the subfiles. Each subfile will be sorted as if it were a separate data set.

   b. Press: CONT

   c. Go to step 7.

5. When "Number of first subfile=?" is displayed:

   a. Enter the number of the first subfile to be sorted.

   b. Press: CONT

6. When "Number of last subfile=?" is displayed:

   a. Enter the number of the last subfile to be sorted. All subfiles between and including the first and last specified will be sorted as though they were separate data sets.

   b. Press: CONT

7. At this point the desired sorting is performed, a hard-copy note of the sort is printed, and the sorted data is stored on file "DATA" of the program medium.
8. When "List data?" is displayed:

   a. Press: YES if a hard-copy listing of the sorted data is desired.

   b. Go to step 1 of the user instructions for LIST.

   or

   a. Press: NO if no listing is desired.

9. When "SELECT ANY KEY" is displayed:

   a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

   b. Go to step 1 of the user instructions for the selected key.

Example

The data was sorted by variable 2. That is, the observations were arranged so that monthly production was in ascending order. Notice that the other variables are carried by the sort operation, so that although the order of the observations has changed, each observation remains intact.

```
     Data set:
                                       SAMPLE
     has been arranged in ascending order according to Variable # 2
     (Subfiles ignored)
       Variable # 1    Variable # 2    Variable # 3    Variable # 4    Variable # 5
          Variable # 6
OBS#
   1
         18.40000      5736.00000        22.00000       146.00000      3110.00000
            30.00000
   2
         21.60000      6116.00000        22.00000       158.00000      3180.00000
            30.00000
   3
         14.90000      6396.00000        21.00000       134.00000      3373.00000
            30.00000
   4
         25.20000      8287.00000        20.00000       171.00000      3293.00000
            30.00000
   5
         22.20000     10768.00000        20.00000       180.00000      3852.00000
            38.00000
   6
         12.50000     11390.00000        20.00000       195.00000      3532.00000
            34.00000
```

7	17.70000	12173.00000	23.00000	191.00000	3366.00000	30.00000
8	18.20000	12257.00000	22.00000	203.00000	3324.00000	30.00000
9	6.90000	12707.00000	20.00000	192.00000	3614.00000	34.00000
10	26.10000	13100.00000	21.00000	196.00000	4345.00000	42.00000
11	27.20000	13108.00000	23.00000	194.00000	4287.00000	42.00000
12	13.30000	13114.00000	19.00000	211.00000	3437.00000	34.00000

Subfiles

Object of Program

This program allows you to specify subfiles or logical groupings of the observations. This may be accomplished by entering the number of observations in each subfile or by entering the observation number of the first observation in each subfile. Names for the subfiles are entered in both cases. A third option allows you to destroy the existing subfile structure.

Special Considerations

Use of Subfiles

Subfiles may be created in order to specify logical groupings of observations. A subfile structure allows you to consider each subfile as a separate data set or to lump all the subfiles together and analyze the overall data set. For example, suppose an investigator wished to measure several variables on 50 trout. He would like to analyze the data separately for each of the three varieties of the trout. He could form three separate data sets and do the individual analyses, then later join the three sets together for the overall analysis. However, since the same variables were measured on each variety of fish, this situation is well-handled by specifying a subfile for each variety. The subfile structure options make it possible to do the analysis by subfile as well as for the overall data set.

Editing and Sorting

Certain operations in the editing and sorting programs may cause observations to move from one subfile to another. To avoid undesired results such as this, it is recommended that subfiles be specified after any editing or sorting has been carried out.

User Instructions

Press: The special function key labeled Subfile.

1. When the subfile characterization options are printed and "Option number=?" is displayed:

   a. Enter 1 if you desire to specify subfiles by entering the number of observations in each subfile.

   b. Press: CONT

   c. Go to step 2.

   or

   a. Enter 2 if you desire to specify subfiles by entering the number of the first observation in each subfile.

   b. Press: CONT

   c. Go to step 2.

   or

a. Enter 3 if you desire to destroy the existing subfiles structure, that is, lump all the data together.

b. Press: CONT

c. Go to step 8.

2. When "Number of subfiles (<= 20) =?" is displayed:

a. Enter the number of subfiles which will be specified.

b. Press: CONT

3. When "Name of subfile #'I' (<= 10 characters) =?" is displayed:

a. Enter the name of the Ith subfile - if may be up to 10 characters long.

b. Press: CONT

4. If the option number you chose was 2, go to step 6.

5. When "Subfile #'I', number of observations =?" is displayed:

a. Enter the number of observations which will be in the Ith subfile.

b. Press: CONT

c. Repeat steps 3 and 5 until all subfiles have been specified.

d. Go to step 7.

6. When "Subfile 'I', number of first observation =?" is displayed:

a. Enter the number of the first observation in the Ith subfile.

b. Press: CONT

c. Repeat steps 3 and 6 until all subfiles have been specified.

7. When "Is above information correct?" is displayed:

a. Press: NO if a mistake has been made while entering the required information.

b. Go to step 2.

or

a. Press: YES if the subfile characterizations are as desired.

8. At this point a hard-copy record of the subfiles will be printed, and the information will be printed on file "DATA" of the program medium.

9. When "SELECT ANY KEY" is displayed:

a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

b. Go to step 1 of the user instructions for the selected key.

Example

The original data (before sorting and recoding) was grouped into two subfiles. The first 12 observations make up the first subfile (Fiscal year '76) and the remaining observations make up the second subfile (Fiscal year '77).

```
Option number = ?
2
Number of subfiles ( <=10 ) = ?
2
Name of Subfile # 1 ( <= 10 characters ) =
?
FY'76
Name of Subfile # 2 ( <= 10 characters ) =
?
FY'77
Subfile # 2 ; number of first observation =
?
13
Is above information correct?
"YES"
Subfile name:    beginning observation--number of observations
 1. FY'76                              1                    12
 2. FY'77                             13                     5

SELECT ANY KEY
```

Name

Object of Program

This program allows you to rename variables and/or to rename subfiles. These names are then stored, along with the data, on the program medium's scratch file ("DATA").

User Instructions

Press: The special function key labeled Name.

1. When "Rename data set?" is displayed:

   a. Press: NO to leave the data set name unchanged.

   b. Go to step 3.

   or

   a. Press: YES to change the data set name.

2. When "Name of data set (<= 80 characters) =?" is displayed:

   a. Enter the project title for the data set.

   b. Press: CONT

3. When the title is printed and "Name variables?" is displayed:

   a. Press: NO if no variables are to be renamed.

   b. Go to step 9.

   or

   a. Press: YES if you desire to rename some variables.

4. When "Name beginning with variable #?" is displayed:

   a. Enter the number of the first variable to be renamed.

   b. Press: CONT

5. When "Name ending with variable #?" is displayed:

   a. Enter the number of the last variable to be renamed.

   b. Press: CONT

6. When "Name of variable #'I' (<= 10 characters)=?" is displayed:

   a. Enter the name of the Ith variable - it may be up to 10 characters long.

   b. Press: CONT

   c. Repeat parts a and b until all specified variables have been renamed.

7. When "Is above information correct?" is displayed:

   a. Press: NO if a mistake has been made while entering variable names.

   b. Go to step 1.

   or

   a. Press: YES if the variable names are as desired.

8. At this point a hard-copy list of all the current variable names will be printed. If there are no subfiles in the data set, go to step 15.

9. When "Name subfiles?" is displayed:

   a. Press: NO if no subfiles are to be renamed.

   b. Go to step 15.

   or

   a. Press: YES if you desire to rename some subfiles.

10. When "Name beginning with subfile #?" is displayed:

    a. Enter the number of the first subfile to be renamed.

    b. Press: CONT

11. When "Name ending with subfile #?" is displayed:

    a. Enter the number of the last subfile to be renamed.

    b. Press: CONT

12. When "Name of subfile #'I' (<= 10 characters)=?" is displayed:

    a. Enter the name of the Ith subfile - it may be up to 10 characters long.

    b. Press: CONT

    c. Repeat parts a and b until all specified subfiles have been named.

13. When "Is above information correct?" is displayed:

    a. Press: NO if a mistake has been made while entering subfile names.

    b. Go to step 10.

14. At this point a hard-copy of the current subfile names is printed.
15. The revised information is stored on file "DATA" of the program medium.
16. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

    b. Go to step 1 of the user instructions for the selected key.

Example

The name of variable four was changed from 'Payroll' to 'Payroll (#)'.

```
Rename data set?
"NO"
Name variables?
"YES"
Name beginning with Variable # ?
4
Name ending with Variable # ?
4
Name of Variable # 4 ( <= 10 characters ) =
?
Payroll(#)
Is above information correct?
"YES"

SAMPLE

CURRENT VARIABLE NAMES
 1. Temp(C)
 2. Production
 3. Days
 4. Payroll(#)
 5. Water Use
 6. Coded Use

Name subfiles?
"NO"
SELECT ANY KEY
```

Store

Object of Program

This program allows you to store the data matrix and related information in a file so that it may be retrieved at a later date for further analysis. The program also allows you to specify the file name.

Special Considerations

Use of Program

The store feature will be useful in two different situations. First, if an investigator has a data set which he may want to analyze further at a later date, he may store it and retrieve it later via the BASIC STATISTICS AND DATA MANIPULATION routine. Secondly, if several people have access to the data input programs, it becomes mandatory that each be able to store his data set in a unique place. Note that if only one person uses the routine on one data set it is unnecessary to use the store feature since the data and related information are kept in "DATA" - the scratch file on the program medium.

Protecting Existing Data

The existence of a file is checked in the program in an attempt to avoid the accidental loss of existing data. Thus, when a file is specified to receive the data, an attempt is made to ensure that you are not accidentally storing the new data in a file which you did not know existed.

User Instructions

Press: The special function key labeled Store.

1. When the program notes are printed and "Name of data file=?" is displayed:

   a. Enter the name of the file in which the data matrix and related information are to be stored - refer to the program note for the proper form of the file name.

   b. Press: CONT

2. When "Is data medium placed in device 'I'?" is displayed:

   a. Press: YES when the medium on which the data and related information are to be stored is in place.

3. When "Does this file already exist?" is displayed:

   a. Press: YES if the file has been previously created (and is of the form specified in the Special Considerations) and you desire to "print over" the information that is currently in the file.

   b. Go to step 4.

   or

   a. Press: NO if the file has not been previously created.

NOTE

This step is an attempt to prevent you from accidentally writing over an existing file. If the response to the question in step 3 is negative and the program finds a file by this name, a note will be displayed and control will pass to step 1.

---

4. At this point, the file will be created if necessary and the data matrix along with the related information will be stored in the specified file. A hard-copy note of the storage will be printed.
5. When "Is program medium replaced in device?" is displayed:
   a. Press:YES when the program medium is in place (or if it was never removed).
6. When "SELECT ANY KEY" is displayed:
   a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.
   b. Go to step 1 of the user instructions for the selected key.

List

Object of Program

This program allows you to obtain a listing of the data matrix. The listing will appear on the device that has been specified for hard-copy in the START routine or in the output unit routine.

User Instructions

Press: The special function key labeled List.

1. When the data listing options are printed and "Option number=?" is displayed:

   a. Enter 1 if you desire to obtain a hard-copy listing of all data - the listing will be by observation.

   b. Press: CONT

   c. Go to step 6.

   or

   a. Enter 2 if you desire to obtain a hard-copy listing of data by variable - it may be all the data or just a set of successive variables.

   b. Press: CONT

   c. Go to step 2.

   or

   a. Enter 3 if you desire to obtain a hard-copy listing of data by observation - the observations are to be successive but need not be the entire data set.

   b. Press: CONT

   c. Go to step 4.

2. When "List beginning with variable #?" is displayed:

   a. Enter the number of the first variable to be included in the list.

   b. Press: CONT

3. When "Ending with Variable #?" is displayed:

   a. Enter the number of the last variable to be included in the list (all variables between the first and last variables specified will be included in the list).

   b. Press: CONT

   c. Go to step 6.

4. When "List beginning with Observation #?" is displayed:

   a. Enter the number of the first observation to be included in the list.

   b. Press: CONT

5. When "Ending with Observation #?" is displayed:

   a. Enter the number of the last observation to be included in the list.

6. At this point the hard-copy listing of the data will be printed.

7. When "SELECT ANY KEY" is displayed:

   a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

   b. Go to step 1 of the user instructions for the selected key.

Join

Object of Program

This program allows you to join or combine two data sets into a single unit. One data set must be in memory and the other data set must have been previously stored by the BASIC STATISTICS AND DATA MANIPULATION program. Two options are available. First, observations may be added together (if both sets have the same number of variables). Second, variables may be added together (if both sets have the same number of observations).

Special Considerations

Adding Observations

Suppose data on six variables was gathered in each of the 52 weeks in 1975, analyzed, and stored on an auxiliary data tape. Suppose the same variables were measured in 1976, analyzed, and stored. If the investigator is interested in lumping the two sets of data together for an overall analysis, he may use the Add Observations option of the joining routine. One set of data must be retrieved via the START routine. Then, after pressing the JOIN key, the second set may be retrieved and the joining carried out. Notice that the variables must be in the same order in the two data sets.

Adding Variables

Suppose an investigator measured five variables on each of 50 mice in an experiment. These were analyzed and stored on disk. Later, he realized that three more variables were of interest. He measured these variables on the mice in the same order as before and analyzed them. All eight variables measured on each mouse could be combined into a single data set via the joining routine

Subfiles

If variables are added, the subfile structure assigned to the resultant data set is the subfile structure of data set #1, that is, the data set that is in the machine prior to the joining operation. If observations are added, the following procedures are employed: 1) If no subfiles exist in either data set, the resultant set has no subfiles. 2) If data set #1 has no subfiles, but data set #2 does, then a subfile named "SET #1" is created which consists of data set #1 and the subfiles of data set #2 remain unchanged. 3) If data set #1 contains subfiles, but data set #2 does not, then a subfile named "SET #2" is created which consists of data set #2 and the subfiles of data set #1 remain unchanged. 4) If both data sets contain subfiles, all of the subfiles of data set #1 are retained and as many subfiles of data set #2 are retained as possible - the upper limit of total subfiles for the resultant set being 20.

User Instructions

Press: The special function key labeled Join.

1. At this point the subprogram "Join" is called. A subprogram is used here for the sole purpose of variable dimensioning.
2. When the joining options are printed and "Option number=?" is displayed:

   a. Enter 1 if you desire to add **variables** from an external data medium to those which are currently in memory.

   b. Press: CONT

   c. Go to step 3.

   or

   a. Enter 2 if you desire to add **observations** from an external data medium to those which are currently in memory.

   b. Press: CONT

3. When the requirements are printed and "OK to continue?" is displayed:

   a. Press: NO if the requirements have not been met.

   b. Go to step 16.

   or

   a. Press: YES if you feel that the requirements have been satisfied.

4. When "Project title for the combined data set (<= 80 characters)=?" is displayed:

   a. Enter up to 80 characters to be used as the name of the joined data set.

   b. Press: CONT

5. When the program note is printed and "File name of data set #2=?" is displayed:

   a. Enter the name of the data file in which the data to be joined to the set already in memory resides. Refer to the program note for the proper form of the file name.

   b. Press: CONT

6. When "Is data set #2 medium placed in device 'I'?" is displayed:

   a. Press: YES when the medium on which data set #2 resides in place.

7. At this point, the number of variables, number of observations, and subfile structures are printed for data sets 1 and 2 as well as for the resultant set. The requirements are checked internally and if they have not been met, control is passed to step 16.

8. When "OK to continue?" is displayed:

   a. Press: NO if, after reviewing the summary, you realize that the resultant data set will not turn out as expected (remember, the subfile structure may be changed at a later time).

   b. Go to step 12.

   or

   a. Press: YES if the resultant data set will be as expected.

9. At this point a summary of the variable names for data sets 1 and 2 as well as for the resulting joined data set are printed.

10. When "OK to continue?" is displayed:

    a. Press: NO if, after reviewing the summary, you desire to abort the joining operation (remember, the variable names may be changed at a later time).

    b. Go to step 12.

    or

    a. Press: YES if the joining operation is to be carried out.

11. At this point the joining operation is performed and a hard-copy summary of the resultant data set is printed.

12. When "Is program medium replaced in device?" is displayed:

    a. Press: YES when the program medium is in place (or if it was never removed).

13. If the joining operation was aborted, go to step 16.

14. At this point the resultant data set and related information are printed on file "DATA" of the program medium.

15. When "List data?" is displayed:

    a. Press: YES if a hard-copy listing of the resultant data set is desired.

    b. Go to step 1 of the user instructions for LIST.

    or

    a. Press: NO if no listing is desired.

16. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.

    b. Go to step 1 of the user instructions for the selected key.

Example

Three observations were entered by keyboard and stored in file "TDATA5". Then the JOIN program added these observations to the end of the original data set to form a new set.

```
Option number = ?
2
OK to continue?
"YES"
Project title for combined data set (<= 80 characters) = ?
SAMPLE
File name of data set #2 = ?
TDATA5:F8
Is data set #2 medium placed in device F8
?
"YES"
OK to continue?
"YES"
OK to continue?
"YES"
Is program medium replaced in device?
"YES"
List data?
"YES"
```

SAMPLE

OBS#	Variable # 1	Variable # 2	Variable # 3	Variable # 4	Variable # 5
1	14.90000	6396.00000	21.00000	134.00000	3373.00000
2	18.40000	5736.00000	22.00000	146.00000	3110.00000
3	21.60000	6116.00000	22.00000	158.00000	3180.00000
4	25.20000	8287.00000	20.00000	171.00000	3293.00000
5	26.30000	13313.00000	25.00000	198.00000	3390.00000
6	27.20000	13108.00000	23.00000	194.00000	4287.00000
7	22.20000	10768.00000	20.00000	180.00000	3852.00000
8	17.70000	12173.00000	23.00000	191.00000	3366.00000
9	12.50000	11390.00000	20.00000	195.00000	3532.00000
10	6.90000	12707.00000	20.00000	192.00000	3614.00000
11	6.40000	15022.00000	22.00000	200.00000	3896.00000
12	13.30000	13114.00000	19.00000	211.00000	3437.00000
13	18.20000	12257.00000	22.00000	203.00000	3324.00000
14	22.80000	13118.00000	22.00000	197.00000	3214.00000
15	26.10000	13100.00000	21.00000	196.00000	4345.00000
16	26.30000	16716.00000	21.00000	205.00000	4936.00000
17	4.20000	14056.00000	22.00000	205.00000	3624.00000
18	25.30000	9315.00000	20.00000	183.00000	3356.00000
19	12.40000	11298.00000	19.00000	203.00000	4205.00000
20	18.60000	14653.00000	21.00000	189.00000	4256.00000

Output Unit

Object of Program

This program allows you to change the device on which the hard-copy output will be printed, or conversely, to specify that no hard-copy is desired.
In lieu of this routine, the hard-copy device may be changed at any time by the following procedure:

Non HP-IB Printer

1. Type: Hc=
2. Key in the select code of the printer desired.
3. Press: EXECUTE
4. Type: Hcbus=999
5. Press: EXECUTE

HP-IB Printer

1. Type: Hc=
2. Key in select code of the printer desired.
3. Press: EXECUTE
4. Type: Hcbus=
5. Key in the bus address of the HP-IB device.
6. Press: EXECUTE

User Instructions

Press: The special function key labeled Output Unit.

1. When the hard-copy options are printed and "Option number=?" is displayed:

   a. Enter 1 if no hard-copy is desired.
   b. Press: CONT
   c. Go to step 5.

   or

   ⋆a. Enter 2 if a hard-copy is desired from the computer's internal printer.
   b. Press: CONT
   c. Go to step 5.

   or

   a. Enter 3 if a hard-copy is desired from an external printer.
   b. Press: CONT

* When using a 9835A, delete option 2 a, b, c and change Enter 3 to Enter 2 if hardcopy is desired from an external printer.

2. When "Are you using an HP-IB Printer?" is displayed:
   (See Appendix for explanation of HP-IB.)

   a. Press: YES if an HP-IB printer is being used.
   b. Go to step 3.

   or

   a. Press: NO if the printer is not HP-IB.
   b. Go to step 4.

3. When "Printer select code, bus address=? [for example, 7,1]" is displayed:

   a. Enter the select code, bus address for the HP-IB printer.
   b. Press: CONT
   c. Go to step 5.

4. When "Printer select code=?" is displayed:

   a. Enter the select code of the printer desired for hard-copy.
   b. Press: CONT

5. When "SELECT ANY KEY" is displayed:

   a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.
   b. Go to step 1 of the user instructions for the selected key.

Basic Statistics

Object of Program

This program computes a variety of summary statistics for data which was entered via the BASIC STATISTICS AND DATA MANIPULATION program set. The statistics may be computed by subfile or for the entire data set (ignoring subfiles). Basic statistics which are computed include: number of observations, number of missing values, sum, mean, variance, standard deviation, coefficient of skewness, coefficient of kurtosis, coefficient of variation, standard error of the mean, and confidence interval of the mean. An option is available to obtain the correlation matrix. Order statistics computed include: the maximum, the minimum, range, and midrange. Additional order statistics which may be obtained include: the median, 25th percentile, 75th percentile, Tukey's middlemeans, and user-specified percentiles.

Special Considerations

Hard-copy Output

If a hard-copy of the statistics is not being made, the program halts occasionally so that you may study the results. In this case, it is necessary only to press 'CONT' to continue program execution.

Additional Order Statistics

If the option to obtain additional order statistics (Tukey's middlemeans and percentiles) is exercised, the data matrix is sorted and the observations of each variable are arranged in ascending order. At the end of the program the original data matrix is re-loaded. Thus, if the program is aborted, that is, if another key is pressed before the re-loading can occur, the data matrix will be in the sorted state. Hence, if the portion of the program used to calculate additional order statistics is accessed, abortion of the program should be discouraged.

Methods and Formulae

Ith Variable

Let N(J) be the number of observations of the Jth variable in the data set or subfile, whichever is applicable. Let D(I,J) be the Jth observations of the Ith variable. The following formulas are computed for the Ith variable.

- Sum: $S(I) = \sum_{J=1}^{N(I)} D(I,J)$

- Mean: $M(I) = \frac{S(I)}{N(I)}$

- Variance: $V(I) = \frac{\sum_{J=1}^{N(I)} (D(I,J))^2 - N(I)(M(I))^2}{N(I)-1}$

- Standard deviation: $Sd(I) = (V(I))^{1/2}$

- Second moment: $M_0(I) = \frac{(N(I)-1)V(I)}{N(I)}$

- Skewness: $Sk(I) = \frac{\sum_{J=1}^{N(I)} (D(I,J))^3 - 3M(I) \sum_{J=1}^{N(I)} (D(I,J))^2 + (2M(I))^3}{(M_0(I))^{3/2} N(I)}$

- Kurtosis:

$$K(I) = \frac{\sum_{J=1}^{N(I)} (D(I,J))^4 - 4M(I) \sum_{J=1}^{N(I)} (D(I,J))^3 + 6(M(I))^2 \sum_{J=1}^{N(I)} (D(I,J))^2 - 3(M(I))^4 N(I)}{(M_0(I))^2 N(I)} - 3$$

t-value

Let C be the confidence coefficient for a confidence interval on the mean. The following operations are used to obtain the desired t-value.

$$P = \frac{1-\frac{c}{100}}{2}$$

$$V = \left(\ln\left(\frac{1}{p^2}\right)\right)^{1/2}$$

$$X = 2.5155174+.802853V+.010328V^2$$

$$Y = 1 + 1.432788V + .189269V^2 + .001308V^3$$

$$Z = V-\frac{X}{Y}$$

$$M = N(J)-1$$

- Then the desired t-value is:

$$T = Z + \frac{Z^3+Z}{4M} + \frac{5Z^5+16Z^3+3Z}{96M^2} + \frac{3Z^7+19Z^5+17Z^3-15Z}{384M^3}$$

$$+ \frac{79Z^9+776Z^7+1482Z^5-1920Z^3-945Z}{92160M^4}$$

- Standard error: $Se(I) = \frac{(V(I))^{1/2}}{(N(I))^{1/2}}$

- Confidence interval on mean: $M(I) \pm T(Se(I))$

- Coefficient of variation: $Cv(I) = \left|\frac{(V(I))^{1/2}}{M(I)}\right| (100)$

- Correlations: Suppose we have the following data matrix:

		OBSERVATION				
		1	2	3	4	5
VARIABLE	1	5	M	3	4	5
	2	6	7	M	6	4
	3	1	3	2	1	1

An M denotes a missing value. When computing the correlation between variables 1 and 2, we discard observations 2 and 3 since variable 1 is missing a data value for observation 2 and variable 2 is missing the data value for observation 3. However, when computing the correlation between variables 1 and 3, we need only discard observation 2. Similarly, the correlation between 2 and 3 is computed by discarding only observation 3. Hence, the correlations may be based on different numbers of observations. An observation is thrown out if and only if a data value from that observation is missing from one of the two variables for which the correlation is being computed. With this in mind, let N(I,J) be the number of observations used to compute the correlation between variables I and J. Then, the correlation is:

$$C(I,J) = \frac{\sum_{K=1}^{N(I,J)} D(I,K)D(J,K) - \frac{\sum_{K=1}^{N(I,J)} D(I,K) \sum_{K=1}^{N(I,J)} D(J,K)}{N(I,J)}}{\left[\sum_{K=1}^{N(I,J)} (D(I,K))^2 - \frac{\left(\sum_{K=1}^{N(I,J)} D(I,K)\right)^2}{N(I,J)}\right]^{\frac{1}{2}} \left[\sum_{K=1}^{N(I,J)} (D(J,K))^2 - \frac{\left(\sum_{K=1}^{N(I,J)} D(I,K)\right)^2}{N(I,J)}\right]^{\frac{1}{2}}}$$

Ranges and Percentiles

Let M(I) be the largest data value of the Ith variable, m(I) be the smallest data value of the Ith variable.

1. Range: $R(I) = M(I) - m(I)$

2. Midrange: $Mr(I) = \frac{M(I)+m(I)}{2}$

3. The percentiles are computed as follows: Let P be the percentile in question. If P*N(I) is an integer, the P(I)=(D(I,P*N(I))+D(1,Q))/2, where Q is the next integer value between P*N(I) and the observation index of the median. If P*N(I) is not an integer, the P(I)=D(I,N(I)*P+Q) where

$$Q = \begin{cases} 1 & \text{if } P \leq 50 \\ -1 & \text{if } P > 50 \end{cases}$$

The median refers to the 50th percentile.

Tukey's Middlemeans

- Midmean: Mm (I) =1/N Σ all observations between (and including, if applicable) 25th and 75th percentile.
- Trimean: Tm(I) =¼ (25th percentile + 2(median) + 75th percentile).
- Midspread: Ms(I) = 75th percentile - 25th percentile.

User Instructions

Press: The special function key labeled Stats.

1. If there is only one variable in the data set, go to step 5.
2. When "Summary statistics on all variables?" is displayed:

   a. Press: YES if summary statistics are desired for all of the variables in the data set.

   b. Go to step 5.

   or

   a. Press: NO if summary statistics are desired for only one variable or a set of successive variables in the data set.

3. When "Number of first variable=?" is displayed:

   a. Enter the number of the first variable for which summary statistics are desired.

   b. Press: CONT

4. When "Number of last variable=?" is displayed:

   a. Enter the number of the last variable for which summary statistics are desired. Summary statistics will be calculated for all variables between and including the first and last specified.

   b. Press: CONT

5. If the data set contains no subfiles, go to step 9.

6. When the subfile structure options are printed and "Option number=?" is displayed:

   a. Enter 1 if you desire to ignore the subfile structure, that is, to lump all the observations together and calculate summary statistics on the data set as a whole.

   b. Press: CONT

   c. Go to step 9.

   or

   a. Enter 2 if you desire to calculate summary statistics for only one subfile or for each of a set of adjacent subfiles; for example, for subfiles 3 through 5. The statistics will be calculated by subfile, that is, calculated as though each subfile were a separate data set.

   b. Press: CONT

   c. Go to step 7.

   or

   a. Enter 3 if you desire to calculate summary statistics for all of the subfiles. The statistics will be calculated as though each subfile were a separate data set.

   b. Press: CONT

   c. Go to step 9.

7. When "Number of first subfile=?" is displayed:

   a. Enter the number of the first subfile for which summary statistics will be calculated.

   b. Press: CONT

8. When "Number of last subfile=?" is displayed:

   a. Enter the number of the last subfile for which summary statistics will be calculated. Statistics will be calculated for all subfiles between and including the first and last subfiles specified as though they were separate data sets.

   b. Press: CONT

9. At this point a heading and basic statistics will be output. If no hard-copy printer has been specified, the program will pause occasionally to allow examination of the output and "Press 'CONT' when ready." will be displayed. To resume execution, press: CONT.

10. When "Confidence coefficient for confidence interval on the mean . . ." is displayed:

    a. Enter the confidence coefficient or confidence level to be used in constructing a confidence interval on the mean; for example, enter 95 for a 95% confidence on the mean.

    b. Press: CONT

11. At this point several more statistics will be output. If the output is to the screen, it will be necessary to press 'CONT' to resume program execution after examining the output.

12. If summary statistics are being computed for a single variable, go to step 15.

13. When "Correlation matrix?" is displayed:

   a. Press: NO if correlations are not desired.

   b. Go to step 15.

   or

   a. Press: YES to obtain a matrix of correlations among the variables.

14. At this point the correlation matrix will be computed and output. If output is to the screen, it will be necessary to press 'CONT' to resume program execution after examining the correlations.

15. At this point several order statistics are computed and output.

16. When "More order statistics (Tukey's middlemeans & percentiles)?" is displayed:

   a. Press: NO if no further order statistics are desired.

   b. Go to step 21.

   or

   a. Press: YES if Tukey's middlemeans (see Methods and Formulae), the 25th, 50th, and 75th percentiles, are desired.

17. At this point the data matrix is sorted by variable. Then, several order statistics are computed and output.

18. When "Other percentiles?" is displayed:

   a. Press: NO if no further percentiles are desired.

   b. Go to step 21.

   or

   a. Press: YES if additional percentiles are desired.

19. When "Press 'NO' when finished - Percentile=?" is displayed:

   a. Enter the percentile which is desired; for example, enter 90 to obtain the 90th percentile. The number entered must be between 1 and 99, inclusive.

   b. Press: CONT

   or

   a. Press: NO if no more percentiles are desired.

   b. Go to step 21.

20. At this point, the desired percentile is computed for each variable under consideration and the results are output. Go to step 19.

21. If summary statistics are being computed for more than one subfile, steps 9-20 are repeated for each subfile.

22. If the additional order statistics were computed for the data set (or any subfile), the original data matrix is re-loaded at this point.
23. When "SELECT ANY KEY?" is displayed:
    a. Press: Any of the keys identified on the BASIC STATISTICS AND DATA MANIPULATION overlay.
    b. Go to the start of the user instructions.

Example

When the basic statistics key (Stats) is pressed, basic statistics and order statistics are printed, first for the entire data set, and then for individual subfiles.

```
*****************************************************************************
*                          SUMMARY STATISTICS                               *
*                             ON DATA SET:                                  *
*                               SAMPLE                                      *
*****************************************************************************
```

(SUBFILES IGNORED)

BASIC STATISTICS

VARIABLE	# OBSERVATIONS	# MISS. VALUES	SUM	MEAN
Temp(C)	20	0	366.50000	18.32500
Production	20	0	232643.00000	11632.15000
Days	20	0	425.00000	21.25000
Payroll	20	0	3751.00000	187.55000
Water Use	20	0	73590.00000	3679.50000

VARIABLE	VARIANCE	STANDARD DEV.	COEF OF SKEWNESS	COEF OF KURTOSIS
Temp(C)	51.93987	7.20693	-.51175	-.87415
Production	9253846.44947	3042.01355	-.64356	-.42042
Days	2.19737	1.48235	.55329	.23443
Payroll	421.62895	20.53361	-1.37170	.95821
Water Use	242672.26316	492.61777	.99654	.11191

VARIABLE	COEF VARIATION	STANDARD ERROR OF THE MEAN	95 % CONFIDENCE INTERVAL ON MEAN LOWER LIMIT	UPPER LIMIT
Temp(C)	39.32841	1.61152	14.95123	21.69877
Production	26.15177	680.21491	10208.09757	13056.20243
Days	6.97578	.33146	20.55607	21.94393
Payroll	10.94834	4.59145	177.93764	197.16236
Water Use	13.38817	110.15268	3448.89172	3910.10828

CORRELATION MATRIX

	Production	Days	Payroll	Water Use
Temp(C)	-.1165685	.2535951	-.1443165	.1428234
Production		.1245049	.8729479	.6347921
Days			-.0272339	-.1524027
Payroll				.4291423

ORDER STATISTICS

VARIABLE	MAXIMUM	MINIMUM	RANGE	MIDRANGE
Temp(C)	27.20000	4.20000	23.00000	15.70000
Production	16716.00000	5736.00000	10980.00000	11226.00000
Days	25.00000	19.00000	6.00000	22.00000
Payroll	211.00000	134.00000	77.00000	172.50000
Water Use	4936.00000	3110.00000	1826.00000	4023.00000

		TUKEY'S HINGES	
VARIABLE	MEDIAN	25-th %-ile	75-th %-ile
Temp(C)	18.50000	12.90000	25.25000
Production	12482.00000	10041.50000	13215.50000
Days	21.00000	20.00000	22.00000
Payroll	194.50000	181.50000	201.50000
Water Use	3484.50000	3340.00000	4050.50000

	TUKEY'S MIDDLEMEANS		
VARIABLE	MIDMEAN	TRIMEAN	MIDSPREAD
Temp(C)	19.29000	18.78750	12.35000
Production	12303.30000	12055.25000	3174.00000
Days	21.20000	21.00000	2.00000
Payroll	193.50000	193.00000	20.00000
Water Use	3544.00000	3589.87500	710.50000

VARIABLE	10-th percentile
Temp(C)	6.65000
Production	6256.00000
Days	19.50000
Payroll	152.00000
Water Use	3197.00000

VARIABLE	90-th percentile
Temp(C)	26.30000
Production	14837.50000
Days	23.00000
Payroll	205.00000
Water Use	4316.00000

```
********************************************************************************
                         SUMMARY STATISTICS                                    *
*                           ON DATA SET:                                       *
*                              SAMPLE                                          *
********************************************************************************
```

Subfile: FY'76

BASIC STATISTICS

VARIABLE	# OBSERVATIONS	# MISS. VALUES	SUM	MEAN
Temp(C)	12	0	212.60000	17.71667
Production	12	0	128130.00000	10677.50000
Days	12	0	257.00000	21.41667
Payroll	12	0	2170.00000	180.83333
Water Use	12	0	42330.00000	3527.50000

VARIABLE	VARIANCE	STANDARD DEV.	COEF OF SKEWNESS	COEF OF KURTOSIS
Temp(C)	50.45242	7.10299	-.24164	-1.10813
Production				
10346097.9091 M7D.5D				
3216.53508	-.47514	-1.24452		
Days	2.99242	1.72986	.53169	-.47442
Payroll	563.60606	23.74039	-.78573	-.59264
Water Use	113867.00000	337.44185	.94917	.15129

VARIABLE	COEF VARIATION	STANDARD ERROR OF THE MEAN	95 % CONFIDENCE INTERVAL ON MEAN LOWER LIMIT	UPPER LIMIT
Temp(C)	40.09212	2.05046	13.20247	22.23086
Production	30.12442	928.53370	8633.28011	12721.71989
Days	8.07718	.49937	20.31728	22.51605
Payroll	13.12833	6.85326	165.74549	195.92118
Water Use	9.56603	97.41107	3313.04397	3741.95603

CORRELATION MATRIX

	Production	Days	Payroll	Water Use
Temp(C)	-.2644635	.4396054	-.1782662	-.0077412
Production		.1294077	.9312213	.6450021
Days			.0040583	.0388568
Payroll				.4892023

ORDER STATISTICS

VARIABLE	MAXIMUM	MINIMUM	RANGE	MIDRANGE
Temp(C)	27.20000	6.40000	20.80000	16.80000
Production	15022.00000	5736.00000	9286.00000	10379.00000
Days	25.00000	19.00000	6.00000	22.00000
Payroll	211.00000	134.00000	77.00000	172.50000
Water Use	4287.00000	3110.00000	1177.00000	3698.50000

		TUKEY'S HINGES	
VARIABLE	MEDIAN	25-th %-ile	75-th %-ile
Temp(C)	18.05000	12.90000	23.70000
Production	11781.50000	7341.50000	13111.00000
Days	21.50000	20.00000	22.50000
Payroll	191.50000	164.50000	196.50000
Water Use	3413.50000	3329.50000	3733.00000

	TUKEY'S MIDDLEMEANS		
VARIABLE	MIDMEAN	TRIMEAN	MIDSPREAD
Temp(C)	18.01667	18.17500	10.80000
Production	11405.50000	11003.87500	5769.50000
Days	21.16667	21.37500	2.50000
Payroll	187.16667	186.00000	32.00000
Water Use	3452.00000	3472.37500	403.50000

VARIABLE	10-th percentile
Temp(C)	6.90000
Production	6116.00000
Days	20.00000
Payroll	146.00000
Water Use	3180.00000

VARIABLE	90-th percentile
Temp(C)	25.20000
Production	13114.00000
Days	23.00000
Payroll	198.00000
Water Use	3852.00000

Subfile: FY'77

BASIC STATISTICS

VARIABLE	# OBSERVATIONS	# MISS. VALUES	SUM	MEAN
Temp(C)	8	0	153.90000	19.23750
Production	8	0	104513.00000	13064.12500
Days	8	0	168.00000	21.00000
Payroll	8	0	1581.00000	197.62500
Water Use	8	0	31260.00000	3907.50000

VARIABLE	VARIANCE	STANDARD DEV.	COEF OF SKEWNESS	COEF OF KURTOSIS
Temp(C)	60.11125	7.75314	-.91030	-.30986
Production	4953615.26714	2225.67187	-.06985	-.38572
Days	1.14286	1.06904	-.75000	-.50000
Payroll	65.41071	8.08769	-.76833	-.75334
Water Use	380730.85714	617.03392	.33651	-1.17044

VARIABLE	COEF VARIATION	STANDARD ERROR OF THE MEAN	95 % CONFIDENCE INTERVAL ON MEAN LOWER LIMIT	UPPER LIMIT
Temp(C)	40.30225	2.74115	12.75397	25.72103
Production	17.03652	786.89384	11202.91630	14925.33370
Days	5.09069	.37796	20.10602	21.89398
Payroll	4.09244	2.85943	190.86170	204.38830
Water Use	15.79102	218.15443	3391.50799	4423.49201

CORRELATION MATRIX

	Production	Days	Payroll	Water Use
Temp(C)	.0040190	-.0844548	-.4503797	.2109508
Production		.4515051	.4876946	.6262209
Days			.2643635	-.3473772
Payroll				.2501814

ORDER STATISTICS

VARIABLE	MAXIMUM	MINIMUM	RANGE	MIDRANGE
Temp(C)	26.30000	4.20000	22.10000	15.25000
Production	16716.00000	9315.00000	7401.00000	13015.50000
Days	22.00000	19.00000	3.00000	20.50000
Payroll	205.00000	183.00000	22.00000	194.00000
Water Use	4936.00000	3214.00000	1722.00000	4075.00000

		TUKEY'S HINGES	
VARIABLE	MEDIAN	25-th %-ile	75-th %-ile
Temp(C)	20.70000	15.30000	25.70000
Production	13109.00000	11777.50000	14354.50000
Days	21.00000	20.50000	22.00000
Payroll	200.00000	192.50000	204.00000
Water Use	3914.50000	3340.00000	4300.50000

	TUKEY'S MIDDLEMEANS		
VARIABLE	MIDMEAN	TRIMEAN	MIDSPREAD
Temp(C)	21.22500	20.60000	10.40000
Production	13132.75000	13087.50000	2577.00000
Days	21.25000	21.12500	1.50000
Payroll	199.75000	199.12500	11.50000
Water Use	3860.25000	3867.37500	960.50000

VARIABLE	10-th percentile
Temp(C)	4.20000
Production	9315.00000
Days	19.00000
Payroll	183.00000
Water Use	3214.00000
VARIABLE	90-th percentile
Temp(C)	26.10000
Production	14653.00000
Days	22.00000
Payroll	205.00000
Water Use	4345.00000

Analysis of Variance

Introduction

The Analysis of Variance package is made up of six analysis routines as well as a number of auxiliary routines that can be used after the analysis of variance (ANOVA or AOV) is completed.

The following analyses are available for balanced data sets -

- **Factorial design** - multiway classification with or without major blocks.
- **Nested design** - includes completely nested, mixed nested and crossed classifications.
- **Split-plot design** - several types in which one or more factors can be in the whole plot.

These three analyses can be used for balanced or unbalanced designs -

- **One-way ANOVA** - completely randomized one-way classification.
- **Two-way ANOVA** (unbalanced) - one or more of the cells can be empty or be unequal in sample size.
- **One-way Analysis of Covariance** - for the completely randomized one-way classification.

For each of the designs in this package, the objective of the routine is to sort out the sources of variability and assign, if possible, responsibility for a portion of the total variability in the data to certain factors in the design.

Input

The first step is to input your data via the Basic Statistics and Data Manipulation routines. Because the data for the AOV programs must be in a very structured format, please read the Basic Statistics and Data Manipulation manual and the section of this manual entitled Data Structures before entering your data. After entering your data, one of the six types of designs is selected and questions will be asked in order to determine the exact design you are using.

Auxiliary Routines

The following routines can be used to complement the analyses performed by the six design routines -

- **Orthogonal Polynomials** - performs a decomposition of the specified sum of squares into linear, quadratic,...,portions. This routine should be used only for factors with quantitative levels.
- **Treatment Contrasts** - performs a comparison on a specified factor. Output includes sum of squares and F ratio.

- **Multiple Comparison Procedure** - can be used to perform one or more of five routines to determine which factor levels represent different population levels.
- **Interaction Plot** - allows you to study the relationship between two or three factors. (Not available from One-way or Covariance routines.)
- **FPROB** - generates right-tailed probability values for the F distribution.

Getting Started

You access the AOV programs via the Basic Statistics and Data Manipulation programs. To get started, follow these steps -

1. Place tape cartridge #1 in the righthand tape drive, :T15.
2. Type LOAD "AUTOST" (or use the Auto start feature of the System 45B).
3. Press: EXECUTE
4. Press: RUN

The program will tell you if you need to switch tapes and insert cartridge #2.

The key labeled "AOV" lets you access AOV routines from Basic Statistics and Data Manipulation. The key labeled Data Manipulation lets you access Basic Statistics and Data Manipulation from AOV. These keys are discussed next.

Accessing AOV

To access AOV after completing the desired operations in Basic Statistics and Data Manipulation press the special function key labeled "AOV".

1. A message will be displayed describing the programs that are on cartridge 1 and those that are on cartridge 2.

   a. If the design you would like to use is on the other cartridge (#2), insert cartridge 2.

   b. Press: CONT

   or

   a. If the desired design is on cartridge 1, press CONT.

2. When "Is the ANALYSIS OF VARIANCE template in place?" is displayed:

   a. Press: YES when the template is in place.

*3. When "SELECT SHIFTED SPECIAL FUNCTION KEY TO SELECT AN APPROPRIATE DESIGN" is displayed:

   a. Press the appropriate key.

   b. Go to the user instructions for that design.

Accessing Basic Statistics and Data Manipulation

While using AOV, pressing the key labeled "Data Manipulation" returns you to the Basic Statistics and Data Manipulation routines.

1. If you are using AOV cartridge 2:

   a. Insert AOV cartridge 1.

   b. Press: CONT

2. When "Is Data Manipulation overlay placed on keys?" is displayed:

   a. Press: YES after placing template on keys.

3. When "Select any key" is displayed:

   a. You may choose any key on the template.

Special Keys

Two special utility keys are provided which allow you to recover from an error and start the program over again, and to specify a new response variable. These keys are discussed next.

Recover

Pressing this key allows you to restart the program you were using, without having to reload it from the cartridge. This key is useful if you feel you have made an error and wish to start over. In general, it is considerably 'safer' to use this key than to press STOP,RUN. After pressing Recover, turn to the beginning of the user instructions for the program you were using.

NOTE

The program will display enough information so that you will know where you are in the entry process.

---

* When using a 9835A, "PRESS" instead of "SELECT SHIFTED" will be displayed.

New Response

Pressing this key allows you to specify a new response variable for the last design chosen. So, even after you have done multiple comparisons (or any other analysis) you may go back to the same design and specify a new response variable without having to answer all of the design questions.

After the key is pressed, a title and description of the last design will be displayed on the CRT.

1. When "Is this the design you want to use?" is displayed:

   a. Press: YES if you wish to use the design specified.

   b. Turn to the User Instructions associated with the desired design.

   c. Questions will start with subfile specifications.

   or

   a. Press: NO if you do not wish to use this design.

   b. Select any key.

Limitations

This program is capable of handling 50 variables with a total of 1500 data values. In addition, there are certain limitations imposed for each program as follows -

- **Factorial** - the product of (levels of A)*(levels of B)*(levels of C)*(levels of D) = size ≤ 500. Also, (number of blocks)*size* (number of observations per cell) ≤ 1500.
- **Nested** - size (as described above) ≤ 500. No blocks are permitted.
- **Split Plot** - Blocks are necessary. Only factors A,B and C are permitted in addition to blocks, and (levels of A)*(levels of B)*(levels of C)*(number of blocks) ≤ 500.
- **One Way** - There can be up to 50 treatments.
- **Two Way (unbalanced)** - At least one cell must have more than one observation. The number of rows (A factor) ≤ 20. The number of columns (B factor) ≤ 20. (number of rows)*(number of columns) ≤ 200.
- **One-way Covariance** - There can be up to 25 treatments.
- **Orthogonal Polynomial** - The polynomial can be up to the tenth degree.
- **Treatment Contrast** - There can be up to 20 levels of one-way means and up to 200 levels of two-way means.
- **Multiple Comparison** - same as for Treatment Contrast.
- **Interaction Plot** - there can be no more than 20 levels of the factor plotted on the X axis, otherwise the plot becomes "messy".

Balanced vs. Unbalanced Designs

To convert from a balanced design to an unbalanced design, you need to use the data manipulation section of the package.

For example, suppose that you have just finished a factorial analysis and now want to use a one-way design on the same data set. To do this -

1. Press the "Data Manipulation" key.
2. Answer any required questions (changing program cartridges if necessary). When "SELECT KEY" is displayed, press "AOV".
3. Make sure that you now have the correct program cartridge for the design you want. Press CONT.

Data Structures

In order to provide for the analysis of six different types of designs the arrangement of the data must be 'presumed' by the program. The material that follows describes the various arrangements within the Basic Statistics and Data Manipulation (BSDM) routines, which are possible for each design. Please read the section dealing with the design which you are considering before attempting to enter your data.

Further information about the designs considered in this package can be found in the Discussion section and in the references.

Factorial Designs

All data to be analyzed with the Analysis of Variance package is entered into memory via the Basic Statistics and Data Manipulation routines. The order in which the data is entered is very important. In general, sampling replications are entered in order, then factors are varied, then blocks are varied. That is, assuming a four-factor design and no sampling replications, the levels of factor D must vary the most rapidly, followed by the levels of C, B, A, and finally the levels of the blocks. Consider an example in which there are two blocks (major replications), two levels of A and three levels of B. Assume for the moment that we do not have any sampling replication and only one response variable. The structure within the Basic Statistics and Data Manipulation (BSDM) program would use only one variable since it is not necessary to store the levels of the factors and blocks when using the (balanced) Factorial program. The structure for this two-way factorial in two blocks would be:

OBS.#	Response Variable 1	Factor B	Factor A	Blocks
1	Y_{111}	B_1	A_1	Block 1
2	Y_{112}	B_2		
3	Y_{113}	B_3		
4	Y_{121}	B_1	A_2	
5	Y_{122}	B_2		
6	Y_{123}	B_3		
7	Y_{211}	B_1	A_1	Block 2
8	Y_{212}	B_2		
9	Y_{213}	B_3		
10	Y_{221}	B_1	A_2	
11	Y_{222}	B_2		
12	Y_{223}	B_3		

NOTE

The levels of Factor B vary most rapidly while the blocks vary the slowest. The Y's represent numerical data which is the only information stored in BSDM. The first subscript indicates the block, the second indicates the level of factor A and the third designates the level of factor B.

You should remember that it is absolutely essential that you arrange your data in this form prior to entering the BSDM program. Of course, if you are careful, there are ways around the apparent limitation suggested above. Consider the following data set which has already been entered via the BSDM program:

OBS#	Variable (i)	Factor V	Factor U	Blocks
1	Y_{111}	V_1	U_1	Block 1
2	Y_{121}	V_2		
3	Y_{112}	V_1	U_2	
4	Y_{122}	V_2		
5	Y_{113}	V_1	U_3	
6	Y_{123}	V_2		
7	Y_{211}	V_1	U_1	Block 2
8	Y_{221}	V_2		
9	Y_{212}	V_1	U_2	
10	Y_{222}	V_2		
11	Y_{213}	V_1	U_3	
12	Y_{223}	V_2		

First of all, note that blocks (major replications) must vary the slowest. We can use this data structure in the Factorial program by telling the program that factor A, the factor which varies slowly, is factor U and has three levels; while factor B is our factor V and has two levels. Hence, independent of the implied subscripts, levels and ordering, we have considerable flexibility in specifying the factors. We must only make sure the Factor A is the factor which varies most slowly while Factor B is the factor which varies most rapidly.

So far we have described how the data must be structured for the major replications and factors. We will now describe the two modes of data arrangement which are permissible for the minor replications (samples). If you have only one sample per treatment combination, there will be no difference between the two modes.

The first mode assumes that the response variable resides in only one of the variables specified in BSDM. Hence any minor replications/samples will have to be entered as subsequent observations in BSDM. For example, suppose we have a factorial with two blocks, two levels of factor A, and three levels of factor B, with two replications (samples) per factorial combination. The data structure with three different response variables might appear as follows:

	Variable				Factor		
OBS#	1=%Ca	2=%Cu	3=%Fe	Sample	B	A	Block
1	X_{11}	X_{21}	X_{31}	1	B_1	A_1	Block 1
2	X_{12}	X_{22}	X_{32}	2			
3	X_{13}	X_{23}	X_{33}	1	B_2		
4	X_{14}	X_{24}	X_{34}	2			
5	X_{15}	X_{25}	X_{35}	1	B_3		
6	X_{16}	X_{26}	X_{36}	2			
7	X_{17}	X_{27}	X_{37}	1	B_1	A_2	

8	X_{18}	X_{28}	X_{38}	2			
9	X_{19}	X_{29}	X_{39}	1	B_2		
10	X_{110}	X_{210}	X_{310}	2			
11	X_{111}	X_{211}	X_{311}	1	B_3		
12	X_{112}	X_{212}	X_{312}	2			Block 2
.							
.							
.							
24	X_{124}	X_{224}	X_{324}	2	B_3	A_2	

The first mode of replicate / sample storage conserves on the use of variables (see Special Considerations for program limitations); however, it does use more observations.

If you have only one response variable in your experiment it may be more efficient to use the second mode for specifying the sampling replications. This mode assumes that each observation in the BSDM program contains all replication values stored one per variable. Hence, the same design described above would appear as follows (here, the subscripts indicate the levels of factor A and factor B, respectively):

	Variables		Factor	Factor	
OBS.	**1=Rep 1**	**2=Rep 2**	**B**	**A**	**Block**
1	X_{11}	X_{21}	B_1	A_1	Block 1
2	X_{12}	X_{22}	B_2		
3	X_{13}	X_{23}	B_3		
4	X_{14}	X_{24}	B_1	A_2	
5	X_{15}	X_{25}	B_2		
6	X_{16}	X_{26}	B_3		
.					
.					
.					

One other example is included without comment. Keep in mind that in our examples we have named the factors A, B, C, and D. As long as your data is arranged in some order with one factor varying the most rapidly within another factor, etc; you can call these factors A, B, C, and D where your factor called A will vary the slowest, etc.

Example (Factorial)--two Blocks, two levels of Factor A, three levels of factor B, two sampling replications:

DATA ENTRY OPTIONS

FORM 1

OBS.#				Variable #1
1	Blk_1	A_1	B_1	Rep_1
2				Rep_2
3			B_2	Rep_1
4				Rep_2
5			B_3	Rep_1
6				Rep_2
7		A_2	B_1	Rep_1
8				Rep_2
9			B_2	Rep_1
10				Rep_2
11			B_3	Rep_1
12				Rep_2
13	Blk_2	A_1	B_1	Rep_1
.				
.				
.				

FORM 2

OBS.#				Variable#1	Variable#2
1	Blk_1	A_1	B_1	Rep_1	Rep_2
2			B_2	Rep_1	Rep_2
3			B_3	Rep_1	Rep_2
4		A_2	B_1	Rep_1	Rep_2
5			B_2	Rep_1	Rep_2
6			B_3	Rep_1	Rep_2
7	Blk_2	A_1	B_1	Rep_1	Rep_2
8			B_2	Rep_1	Rep_2
9			B_3	Rep_1	Rep_2
10		A_2	B_1	Rep_1	Rep_2
11			B_2	Rep_1	Rep_2
12			B_3	Rep_1	Rep_2

The order of the observations must be as shown above to get the correct results. In general, the levels of blocks will vary slower than levels of factor A, B, C, D and replicates within cells vary the fastest.

Nested Design

The form of the data structure for the nested or mixed design is quite similar to that previously described for the Factorial Designs. As far as the program is concerned, the nested design is considered to be in a factorial arrangement. The program will calculate the sum of squares, etc., as if the design were a factorial design and then pool the appropriate terms to form the nested or mixed design which you specified.

As you may have already noted, the design must be balanced. This means that if factor C is nested within factor A and is denoted as C(A), then there must be exactly the same number of levels of factor C within each level of factor A. You may wish to refer to the Discussion section to familiarize yourself with the design arrangements for a nested design as compared to a factorial design.

Perhaps an example of a completely nested design structure would be helpful at this time. Suppose that within each of five sections of land we select two lakes at random. From each lake assume that three random positions in the lake are chosen at which we select two samples. Suppose further that the samples are each divided into two beakers and are analyzed separately. Assume that three responses are measured: Y_1=Var.1=ppm lead, Y_2=Var.2=ppm zinc, and Y_3=Var.3=ppm copper.

In this experiment, we will designate the five land sections as the levels of factor A, the various lakes as levels of factor B, and the position as levels of factor C. Notice that factor B is nested in factor A, and that factor C is nested within factor B. These relationships are commonly denoted by B(A) and C(B) respectively.

For the first form of data arrangement, the two samples per position in the lake will be shown as stored in subsequent observations (down) rather than in an additional variable (across). A dash (−) indicates a numerical value which would be entered in BSDM.

Form 1

Obs#	Var1=Y_1	Var2=Y_2	Var3=Y_3	Sample	Position	Lake	Section
1	-	-	-	1	P_1	L_1	Sec 1
2	-	-	-	2			
3	-	-	-	1	P_2		
4	-	-	-	2			
5	-	-	-	1	P_3		
6	-	-	-	2			
7	-	-	-	1	$P_1=P_4$*	L_2	
8	-	-	-	2			
9	-	-	-	1	$P_2=P_5$		
10	-	-	-	2			
11	-	-	-	1	$P_3=P_6$		
12	-	-	-	2			
.				.	.	.	.
.				.	.	.	.
.				.	.	.	.
60	-	-	-	2	$P_3=P_{30}$	$L_2=L_{10}$**	Sec_5

*Within each lake the 'first' position P_1 has no relationship with the 'first' position in another lake; hence we have a total of thirty different lake positions.

**Since each section has two lakes selected from it, there are a total of ten lakes studied in this project.

The other form of data entry for this nested design would use twice as many variables since each sample would be included as another variable rather than another observation. Hence the last row would look like:

	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
Obs#	Var1=Y_1	Var2=Y_1	Var3=Y_2	Var4=Y_2	Var5=Y_3	Var6=Y_3
30	-	-	-	-	-	-

With a little practice you will find that it is quite easy to structure your data so that the Nested Analysis will correctly recognize your data set.

Mixed designs must be entered via the BSDM routines in a similar manner. Keep in mind that whichever factor you call D must have its levels varying more rapidly than factor C which in turn varies faster than factor B. The levels of factor A will change only after each level of factor B have appeared once.

NOTE

BLOCKS as described in the Factorial Design are not considered for the Nested Design. That is, you will not be asked any questions concerning blocks (major replications) of this design.

Split-Plot Design

In terms of the data structure in the BSDM routine, it is immaterial whether one is using a Split-Plot Design or a Factorial Design. Both designs are the same in terms of the data arrangement in BSDM. Examples representing the two modes of data arrangement for the minor replications (samples) will be shown below. Consider a split-plot experiment in which the pull-off force necessary to remove boxes from a tape is to be studied (see Hicks pp 219-222, 226). Two complete replications (blocks) of the following experiment were performed. Three long strips of tape with boxes attached were chosen to represent three different methods of attaching the boxes to the strips. A chamber was used to study the effects of three humidity levels (50, 70, and 90%) on the pulling force of three boxes. The experimental procedure called for randomly choosing one of the three humidity levels and adjusting the chamber to maintain that level. Two portions of each of the three strips were placed in the chamber for a specified period of time. The pull-force was then measured for each of the six portions of strip. Subsequently, one of the two remaining levels of humidity was randomly chosen and the process was repeated. Finally, the last level of humidity was maintained in the chamber. Upon completion of the first three humidities times three strips times two samples = 18 measurements, the entire process was repeated again in a random fashion.

The reason that this is a split-plot design and not a factorial is because of the ordering of the measurements of pull force. Since it was not deemed possible to randomly investigate the effects of humidity and strip type on the pull force response, we have a restricted randomization of the split-plot type.

The two forms for specifying the sample replications are shown below. Note how the factor names A and B have been assigned to the factors in this experiment and how that corresponds to the data arrangement as shown. Only one response variable is necessary for this design.

FORM1

OBS#	Y=pull force Variable 1	Sample	B Humidity	A Strip	Block
1	-	1	50%	S1	B1
2	-	2			
3	-	1	70%		
4	-	2			
5	-	1	90%		
6	-	2			
7	-	1	50%	S2	
8	-	2			
9	-	1	70%		
10	-	2			
11	-	1	90%		
12	-	2			
13	-	1	50%	S3	
14	-	2			
15	-	1	70%		
16	-	2			
17	-	1	90%		
18	-	2			
19	-	1	50%	S1	B2
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
36	.	2	.	.	.

In this experiment we would specify two blocks (major replications). Factor A (strips) has three levels, factor B (humidity) has three levels, and there are two samples for mode 1 (all samples wihin the same variable). Later, in the Split-Plot Design program, we would specify that factor B (humidity) is the whole plot while factor A (strips) is the subplot. As the experiment is described above, the humidity factor (B) would be in the whole plot even though it does not vary as fast as the strip factor (A). We could have entered our data in a manner which would have had the levels of humidity varying the slowest. Then we would identify humidity as factor A.

The second mode of sample specification for this example would require two variables, say variable one and variable two.

FORM2

OBS#	Y=pull force Var 1=Sample 1	Var 2=Sample 2	B Humidity	A Strip	Block
1	-	-	50%	S_1	B_1
2	-	-	70%		
3	-	-	90%		
4	-	-	50%	S_2	
5	-	-	70%		
6	-	-	90%		
7	-	-	50%	S_3	
8	-	-	70%		
9	-	-	90%		
10	-	-	50%	S_1	B_2
11	-	-	70%		
12	-	-	90%		
13	-	-	50%	S_2	
14	-	-	70%		
15	-	-	90%		
16	-	-	50%	S_3	
17	-	-	70%		
18	-	-	90%		

One-Way Design

The one-way design, or one-way classification as it is sometimes called, has three possible forms of data organization or structures in BSDM. These three forms are identical to the forms for the ONE-WAY ANALYSIS OF COVARIANCE except that the covariance analysis will expect both a response variable, Y, and a covariate, X, to be specified while the ONE-WAY DESIGN expects only the response variable Y.

The first mode of data organization for the one-way classification uses t variables in BSDM to specify the t treatments in this design. Consider an experiment in which four types of "mums" were investigated in a greenhouse experiment. Suppose two responses were measured: diameter (Y_1) and plant height (Y_2). The data was collected in two separate years (subfiles) with approximately five pots per variety. One possible organization of this data is as follows:

Mode 1 Example

	Variable	1	2	3	4	5	6	7	8
	Response	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
	Treatment/Variety	Type 1		Type 2		Type 3		Type 4	
	OBS#								
Subfile 1975	1	-	-	-	-	-	-	-	-
	2	-	-	-	-	-	-	-	-
	3	-	-	-	-	-	-	-	-
	4	-	-	MV	MV	-	-	-	-
	5	-	-	MV	MV	MV	-	MV	MV
Subfile 1976	6	-	-	-	-	-	-	-	-
	7	-	-	-	-	MV	-	-	-
	8	MV	MV	-	-	-	-	-	-
	9	MV	MV	-	-	-	MV	-	-
	10	MV	MV	-	-	-	-	-	-
	11	MV	MV	MV	MV	-	-	-	-

Here, a dash (−) indicates a numerical value is present, and MV indicates that a missing value is assigned to this position.

NOTE

The arrangement shown above has provisions for missing values to accommodate the various number of pots per treatment (variety). The two subfiles do not have the same number of pots per treatment. The MV operation must be used to 'square-off' the sample sizes for each variable.

You would tell the program that variables one, three, five, and seven represent the four treatments for the first response (diameter). You would then specify the subfile number. The program would then assume that the sample size is five if subfile one is specified and six if subfile two is specified. If subfiles are to be ignored, then a sample size of 11 would be assumed. Of course all calculations within the program would check for missing values (MV) and delete those values from the calculations. Subsequent to the analysis on the first response, Y_1, you may remain within this subfile and specify another response, say Y_2. Finally, you may select another subfile and/or variables for further analysis.

The second mode for possible data organization within the BSDM structure uses only one variable for each response. Within this response variable, the treatment observations are assumed to be contiguous. You specify the number of observations in each treatment including any missing values. The program assumes that the first observation in the first treatment is observation number one if the first subfile is chosen or subfiles are ignored, or the first observation within the specified subfile. Thereafter, the subfile is partitioned into t nonoverlapping but connected intervals - one corresponding to each treatment. Hence, for the example with four treatments and two response variables, one possible arrangement might be:

Mode 2 EXAMPLE

	OBS#	Variable 1 Y_1	Variable 2 Y_2	Treatment# (Variety#)
SUBFILE 1 1975	1	-	-	1
	2	-	-	
	3	-	-	
	4	-	-	
	5	-	-	
	6	-	-	2
	7	-	-	
	8	-	-	
	9	-	-	3
	10	-	-	
	11	-	-	
	12	-	-	
	13	MV	-	
	14	-	-	4
	15	-	-	
	16	-	-	
	17	-	-	
SUBFILE 2 1976	18	-	-	1
	19	-	-	
	20	-	-	2
	21	-	-	
	22	-	-	
	23	-	-	
	24	-	-	
	25	-	-	3
	26	MV	-	
	27	-	-	
	28	-	MV	
	29	-	-	
	30	-	-	
	31	-	-	4
	32	-	-	
	33	-	-	
	34	-	-	
	35	-	-	
	36	-	-	

NOTE

The sample sizes for the first subfile of each variable would be five, three, five, and four, respectively. For subfile two, the sample sizes would be two, five, five, and six. Of interest is the comparison between the number of data storage positions needed for the two modes of arrangement. For mode 1, the number of positions required would be 11 observations times 8 variables = 88. For the second mode, the number required is 36 observations times 2 variables = 72. In many cases, if there are several missing values you may conserve available memory locations by using the second mode of arrangement.

The third mode of data entry allows for treatments which are not necessarily connected within one variable. Each treatment is composed of a contiguous set of observations. Since this mode of data arrangement may choose treatment groups throughout the data set, it is not possible or necessary to specify subfiles. The arrangement of the data is similar to the arrangement described for method 2, however it is possible to have "gaps" or "holes" in the data set.

Consider the example described above. Suppose it is desired to compare 1975 variety #2 with the 1976 variety #2 for both responses (Y_1 and Y_2). Please refer to the Mode 2 Example and note that we would need to compare observations 6, 7, and 8 with observations 20, 21, 22, 23, and 24. The first three specified observations are from variety #2 in subfile one which is the 1975 data set and the other five values are from variety #2 in subfile two which is the 1976 data set.

Note that although this mode of data arrangement is quite similar to Mode 2, it does provide for more freedom on the part of the data analyst in terms of which treatments are to be used.

Two-Way (Unbalanced) Design

The unbalanced nature of this design makes it more complicated in terms of the data arrangement. It will not be possible to assume that the order of input is completely specified by factor names such as factor A and factor B. This is because it is possible to have not only different numbers of minor replication (samples) within each treatment combination (levels of factor A and factor B), but also to have one or more cells completely missing. Of course, the absence of certain cells is not a desirable characteristic of any factorial experiment; however, there are certain situations in which missing cells naturally occur.

Therefore it is necessary for the BSDM data structure to provide for proper identification of the row and column levels (factors A and B) as well as the particular sample number within that cell. Two methods of specification are permitted for this type of design. The first "data storage type" assumes that you will use three BSDM variables to specify the response variable and factor levels. One variable will be used to store the particular response to be analyzed at this time. One variable will be used for each of the two factors A and B. It is not necessary to use a variable to specify the sample or observation number; however, you may wish to do so in order to completely identify each observation.

Please note that the levels of factors A and B must be the integers 1, 2,...up to the number of levels of each factor. Hence, if factor A has three levels 70, 80, and 120, you would store these three levels in a variable as 1, 2, and 3 rather than 70, 80, and 120. The purpose of this restriction is to conserve data storage allocation. Within the program you will be able to specify the actual levels of the variables when this is necessary for the computation.

As an example of the first data storage type, suppose you have factors of time and temperature involved in an experiment which is designed to study the effects of these two factors on the yield (Y) of a chemical process. Suppose you had used three time settings of 4, 5, and 7.5 hours and three temperature settings of 110, 115, 120° F. Assume that, for one reason or another, from two to five samples were run at each treatment combination (temperature and time condition). Further, let us assume that at the highest temperature and time condition, it was impossible to finish the experimental process. Thus, we can consider this "cell" as missing. Assume two responses Y_1 and Y_2 were measured on almost all samples. One way to enter this data set in the BSDM program is as follows:

Mode 1 Example

BSDM Variable Number

Obs #	**1 Y_1**	**2 Y_2**	**3 B Levels**	**4 A Levels**	**Sample**	**A Temp**	**B Time**
1	MV	-	1	1	1	110°	4 hrs.
2	-	-	1	1	2		
3	-	-	1	2	1	115°	
4	-	-	1	2	2		
5	-	-	1	2	3		
6	-	-	1	3	1	120°	
7	-	-	1	3	2		
8	-	-	1	3	3		
9	-	-	1	3	4		
10	-	-	2	1	1	110°	5 hrs.
11	-	-	2	1	2		
12	-	-	2	1	3		
13	-	-	2	2	1	115°	
14	-	-	2	2	2		
15	-	-	2	3	1	120°	
16	-	-	2	3	2		
17	-	-	2	3	3		
18	-	-	2	3	4		
19	-	MV	2	3	5		
20	-	-	3	1	1	110°	7.5 hrs.
21	-	-	3	1	2		
22	-	-	3	1	3		
23	-	-	3	2	1	115°	
24	-	-	3	2	2		
25	-	-	3	2	3		
26	MV	MV	3	3	1	120°	

Notes:

1. Observation number 26 is included to let the program know that the cell with temp=120, time=7.5 is missing in both responses.
2. Both observation #1 and #19 have one and only one missing response.
3. Although we have shown the 26 observations in a systematic arrangement, this is not necessary except for your own information.
4. The specification of variable numbers in the analysis will identify which factor it should consider as rows (factor A) and which it should consider as columns (factor B).

The second data storage mode allows you to conserve on variables by using only one variable to identify both row and column levels. The levels are "packed" into four digits as xxyy, where xx identifies the row level and yy identifies the column level. Consider the example described above. Using the packed form of storage we will need to allocate at least three variables in the BSDM routine. One variable is needed for each response and one for the 'packed' row/column identification. You may wish to use another variable to identify the sample numbers or you might wish to use the 'space' after the row/column specification. For example, suppose for the third row and second column you wish to identify the observation by the index 74. The packed version would be 0302.74. The program will use only the first four digits 0302 to identify the row and column numbers. Up to 6 digits may be input after the decimal point for identification purposes.

The example described above may be entered via the BSDM routine as follows (for the first ten and the last three observations):

Mode 2 Example

BSDM Variable Number

Obs #	1 Y_1	2 Y_2	3 ID xxyy	Obs#	A Temp	B Time
1	MV	-	0101	1	110°	4 hrs.
2	-	-	0101	2		
3	-	-	0102	1	115°	
4	-	-	0102	2		
5	-	-	0102	3		
6	-	-	0103	1	120°	
7	-	-	0103	2		
8	-	-	0103	3		
9	-	-	0103	4		
10	-	-	0201	1		
.	-	-	-	-		
.	-	-	-	-		
.	-	-	-	-		
24	-	-	0302	2		
25	-	-	0302	3		
26	MV	MV	0303	1	120°	7.5 hrs

One-Way Covariance

The three forms of data arrangement for the one-way analysis of covariance are the same as the one-way design except that both a response variable (Y) and a covariate (X) must be specified. Hence, for the example previously described for mode 1 of the one-way design you would need to specify 12 variables of the BSDM data set and specify a covariate for each treatment set. If different covariates are to be used with the two response variables, then you would need 16 variables. One possible ordering of these variables and treatments for the ith observation is as follows:

	Type 1			Type 2			Type 3			Type 4		
Variable#	1	2	3	4	5	6	7	8	9	10	11	12
	X	Y_1	Y_2	X	Y_1	Y_2	X	Y_1	Y_2	X	Y_1	Y_2

For both mode 2 and 3, you would need to specify one additional variable number as the covariate for each dependent variable. Of course the response variables may use the same covariate in the analysis.

Discussion

General

The analysis of variance (AOV) technique can be used in many data analysis situations where it is desired to characterize the sources of variation in a "planned" experiment. The essential feature of AOV is that the total variation of the numbers (data) is uniquely decomposed into separate parts. For example, suppose we have run an experiment in which we used four varieties of corn and three row spacings. We repeated this experimental set-up five times (on five fields). We can then break the total variation down into five components as indicated below:

AOV

Source	DF	SS	MS	F
Total	5*4*3−1=59	SS_T		
Fields (or Blocks)	5−1=4	SS_B	MS_B	$F_1=MS_B/MS_E$
Varieties	4−1=3	SS_V	MS_V	$F_2=MS_V/MS_E$
Row Spacings	3−1=2	SS_R	MS_R	$F_3=MS_R/MS_E$
Var. X Row	3*2=6	SS_{VR}	MS_{VR}	$F_4=MS_{VR}/MS_E$
Error	44	SS_E	MS_E	

In order to more fully develop our understanding of the usefulness of AOV, let us discuss how one might use such a table. Starting with the first column, we see the decomposition of the total variation into its five components. The next column shows the allocation of the so-called degrees of freedom (see references). Notice that the degrees of freedom components add up to the degrees of freedom associated with the total sum of squares. For the total source of variation, the degrees of freedom will be the total number of observations in the experiment minus one. The SS(sum of squares) column shows the breakdown of the total sum of squares for the experiment into the various components. One could prove algebraically that $SS_T = SS_B + SS_V + SS_R + SS_{VR} + SS_E$ and likewise for the degrees of freedom. The MS (mean square) column is obtained by taking SS/DF. This reflects an "average" variation due to each of the sources.

The last column is the F-ratio or testing column. Generally, we are testing the hypothesis that there is "nothing" happening in the experiment versus the expected hypothesis that something "worthwhile" is occurring. If nothing is happening, then all mean sources of variation should be of the same magnitude as the error mean square. The F-ratio is a statistical test to see if the mean square for the source of variation in question is significantly bigger than the error mean square. If it is, we can conclude that there is a "real" effect. For example, suppose that F_2 is quite large. We would then be able to conclude that the population variety means are not all the same. That is, at least one of the variety means differs significantly from the others.

How big do the F values have to be? That depends on the degrees of freedom associated with the numerator MS and the degrees of freedom associated with the denominator (error) MS. The computed F values may be compared with tabled values to find out if they are significant at the .10, .05, .01, or .005 level, or, with this program, you can actually compute the level of significance. The program will automatically calculate the Prob[F > F calculated] for a factorial AOV. For nested or partially nested AOV, the user may elect to use the FPROB key to find the probability levels.

Factorial Versus Nested Models

Many researchers have difficulty differentiating between a factorial model and a nested model for AOV. A brief example may be of some help. In a three-way factorial model, for example, the levels of factor B are the same over all levels of factors A and C. Suppose factor A is three temperature settings, factor B is two pressure settings and factor C is four different laboratories. In a factorial model, we would assume that each of the six (three temperature * two pressure) combinations had been studied at each of the four laboratories. In a nested AOV with factor C nested in A and B, we might assume that the same six combinations were run; however, for each of the six combinations, four different laboratories (greenhouses, plants, fields, classrooms, etc.) were used. Hence, a total of 24 laboratories were used instead of just four. Assuming just one observation per laboratory and experimental combination, the AOV table for the factorial would be:

Factorial AOV Example

Source	DF	SS	MS
Total	3*2*4−1=23	SS Total	
Temperature	3 − 1 = 2	SS_T	MS_T
Pressure	2 − 1 = 1	SS_P	MS_P
Temp x Pres	2 * 1 = 2	SS_{TP}	MS_{TP}
Laboratories	4 − 1 = 3	SS_L	MS_L
Temp x Lab	2 * 3 = 6	SS_{TL}	MS_{TL}
Pres x Lab	1 * 3 = 3	SS_{PL}	MS_{PL}
Temp x Pres x Lab	2*1*3 = 6	SS_{TPL}	MS_{TPL}

However, for the nested model described above, the AOV table would be:

Nested AOV Example

Source	DF	SS	MS
Total	23	SS_{Total}	
Temperature	3 − 1 = 2	SS_T	MS_T
Pressure	2 − 1 = 1	SS_P	MS_P
Temp x Pres.	2 * 1 = 2	SS_{TP}	MS_{TP}
Lab (temp x pres)	(4−1)*3*2=18	$SS_{L(TP)}$	$MS_{L(TP)}$

Notice that the AOV tables are somewhat different. Actually, the $SS_{L(TP)}$ can be obtained (and is in the program) from the first AOV table by noting that $SS_{L(TP)} = SS_L + SS_{TL} + SS_{PL} + SS_{TPL}$. Generally, in nested or partially nested AOV's, the nested factor is considered to be a random effect.

Partially Nested vs. Nested Models

Consider a laboratory experiment involving mice in which three levels of some drug (factor A) are to be investigated. Seven mice (factor B) are used for each drug level and the response variable is determined on four days (factor C). One model which might be used for the analysis would be three levels of factor A; seven levels of factor B nested on factor A; and four levels of factor C. The AOV table would be:

AOV

Source	DF	SS	MS
Total	83	SS_{TOTAL}	
Drug	2	SS_D	MS_D ←
Mice(Drug)	18	$SS_{M(D)}$	$MS_{M(D)}$
Days	3	SS_T	MS_T ←
Drug x Time	6	SS_{DT}	MS_{DT} ←
Time x Mice(Drug)	54	$SS_{TM(D)}$	$MS_{TM(D)}$

This type of design is sometimes called a repeated measurements design. It is also a partially nested design because factor C is crossed both with factor A and the nested factor B. As is indicated by the arrows in the AOV table, at least two different "error" terms are used for studying the significance in this model. It should be noted that it is necessary to have exactly the same number of subjects within each level of factor A in order to use the analysis in this package.

Two-Factor AOV Structure

The analysis of variance is a method of decomposing the sum of squared deviations of the observations about the overall mean $[\Sigma(y_{ijk} - \bar{y}_{...})^2]$ into various sources. For a two-factor design, we may show sources of variation due to the row effect (A), the column effect (B), the row-by-column interaction effect (AB) and the within error effect (ERROR). For example, consider an experiment in which we have four levels of temperature (100, 150, 175, 200°C) and three levels of pressure (5, 10, 15 psi) with several determinations of the chemical yield (y) for each combination of temperature (ROWS) and pressure (COLUMNS). One possible arrangement of the data might be as shown below:

			Pressure	
		5	10	15
Temperature		Column 1	Column 2	Column 3
100	Row 1	$y_{111},\ldots,y_{11n11}$	$y_{121},\ldots,y_{12n12}$	$y_{131},\ldots,y_{13n13}$
150	Row 2	.	.	.
175	Row 3	.	.	.
200	Row 4	$y_{411},\ldots,y_{41n41}$	$y_{421},\ldots,y_{42n42}$	$y_{431},\ldots,y_{43n43}$

Each y_{ijk} stands for the numerical value of the chemical yield in percent. The subscript i refers to the row designator, the j for the column designator, and the k for the observation number in the i,jth cell. Notice that the n_{ij} are not necessarily all equal, nor is it necessary that n_{ij} be $>= 1$. If the n_{ij} are all equal, the analysis of variance involves the usual summing and summing of squares, a task which could be performed by hand calculators. When the n_{ij} are not all equal, the exact analysis is quite complicated.

Note that the table which we have described above does not show how the experiment was actually run. According to good statistical practice the order of running the experiment should be in a random fashion. That is, conceptually, all of the possible sequences should be equally likely and the experimenter should choose one sequence at random.

Reasons for Unbalanced Designs

Unbalanced two-factor designs might arise in at least three ways. First, the design could have been planned as a balanced design (all n_{ij} equal). However, several observations may be lost due to death of a subject, etc. This often happens in research even though experimenters use good experimental techniques. Second, because of the nature of the variability of one response (or some other reason), the experimenter may have set up the design with an unequal number of observations in the cells. For example, suppose that one of the row levels is really a control or standard dose. It may be a common practice to use fewer observations on the control than the other drugs (other "levels" of the row factor). A third possibility is that certain combinations of the row and column levels might yield results which are impossible to monitor in an experiment. This might happen if in the experiment described above, the highest temperature level (200°C) and the highest pressure level (15 psi) proved to be "too much" for the chemical process. In general, of course, it is not a good procedure to design two-factor experiments in which certain levels of the factors cannot be included in the experiment.

Approximate Analyses for Two-Factor Experiments

If each cell (row-column combination) has at least one observation and the number of observations in each cell is approximately the same, the method of unweighted means is sometimes used. Essentially, in this analysis, the cell means are subjected to the usual two-way AOV with one observation per cell, and the within error term is added to the table after adjustment. (See Bancroft, reference 1, p. 35.) This approximate analysis will probably allow you to draw accurate conclusions for most sets of data.

One reason why we might use this type of analysis is because the "exact" analysis is quite complicated. The complexity of the analysis is related to the fact that the calculations which must be performed do not just involve the usual summing and summing of squared values. In short, the exact analysis is a "messy" problem.

Unbalanced Two-Way AOV - "Exact" Solutions

As described more completely in reference 1, Chapter 1, the solution involves rather messy notation. We shall avoid the notational problems by describing, in words, the procedures that you should use in interpreting the AOV tables, rather than describing the computing procedures which were used.

Once again, the idea of the AOV is to separate out the various sources of variation from an observable set of data. In the balanced two-factor design, the analysis of variance table might be written as follows:

AOV

Source	df	Sum of Squares	Mean Squares
Total	N−1	TSS	
Rows	R−1	RSS	RSS ÷ (R−1)
Columns	C−1	CSS	CSS ÷ (C−1)
RxC Interaction	(R−1)(C−1)	ISS	ISS ÷ (R−1)(C−1)
Residual	N−RC	ESS	ESS ÷ (N−RC)

In this table, R equals the number of rows, C equals the number of columns, and N equals the number of observed y's. The computations which are involved in obtaining the Sum of Squares column will not be described. Suffice it to say that in each case the individual observations or the means are compared to the overall mean.

As a brief review, let us examine that AOV procedure. According to the AOV procedure, we are trying to determine if the source of variation for rows, columns, and/or the interaction is significantly bigger than the error source of variation. This is done by calculating certain ratios of means squares--the so-called F-ratios. Under the assumption of no differences among the row population means (i.e., levels of temperature), the mean square (MS) for rows should be of the same magnitude as the MS for the error. In a similar fashion, the source of variation for columns and interaction can also be tested.

For balanced sets of data, that is where the subclass frequencies are all the same, the decomposition of the sources of variation for a two-factor design is orthogonal. This means that every SS and MS in the table represents the source of variation as indicated in that row. When we have an unbalanced design, the table is not as easy to interpret.

In order to understand the output provided by this program, we will use the hypothetical experiment described earlier. Suppose that the table of n_{ij}, the frequency counts for the twelve row-column cells is as follows:

		Pressure			
		5	10	15	
Temperature	100	5	4	5	(N=54)
	150	5	5	5	
	175	5	5	4	
	200	4	3	4	

Ordinarily we would ask the investigator to use equal n_{ij}; however, there might be perfectly good reasons why this was not possible.

Preliminary AOV Tables

The next output from this program is the Preliminary AOV tables. The first table has the general form:

Preliminary AOV

Source	DF	SS	MS	F-ratio
Total	N−1 = 53	SS_T		
Subclass*	RC−1 = 11	SS_S	MS_S	MS_S/MS_E
ERROR	N− RC = 42	SS_E	MS_E	

*Rows + Columns + Interaction

The decomposition in this table looks as if we have twelve individual treatments rather than four temperature and three pressure combinations. If the F-ratio is large (and the F-Prob is small), say less than about .05, we can conclude that not all twelve population means are the same. The second table has a further decomposition of the subclass source into main effect differences and interaction differences.

Interaction Preliminary AOV

Source	DF	SS	MS	F-Ratio
Total	N−1 = 53	SS_T		
Main Effects*	R+C−2 = 5	SS_M	MS_M	MS_M/MS_E
Interaction**	(R−1)(C−1) = 6	SS_I	MS_I	MS_I/MS_E
Error	N−RC = 42	SS_E	MS_E	

* Row + Column
**RxC

This table helps us determine if there is interaction in our two-way design. This is important because it may help us decide which analysis to use next, that is, which of the FINAL AOV's we should choose (see Bancroft).

If one or more cells are empty, the method of fitting constants must be used for the final analysis. For the method of fitting constants, we assume no interaction is present in the model. Hence, if either one $n_{ij} = 0$ and/or interactions are assumed to be absent in the population, we should use the METHOD OF FITTING CONSTANTS FINAL AOV. If interaction between the row and column factors is expected to be present in the population and all $n_{ij} >= 1$, the METHOD OF SQUARED MEANS should be used.

If you are uncertain whether or not interactions are present, your interpretation of the output of the PRELIMINARY AOV table for interactions may help you decide. If the F-PROB for the interaction F-ratio is small enough, we might conclude that interaction is present. (Bancroft, reference 1, suggests that if F-PROB < .25, one should use the method of squared means.)

Interpreting the Method of Fitting Constants AOV

Since this method assumes that the model is of the form Y = A + B * (ROW LEVELS) + C * (COLUMN LEVELS) + ERROR, what remains to be tested by this method is if the row levels (means) differ significantly from each other and if the column levels (means) differ significantly from each other. The calculations involve (see page 16, Bancroft) finding the solution to a set of least-squares equations. As we discussed above, when all n_{ij} are equal, the sum of squares due to rows is orthogonal to the sum of squares for columns. However, when the n_{ij} are not all equal, by using the method of fitting constants, the program will construct the following table:

Source	DF	SS	MS	F-Ratio
Total	$N-1 = 53$	SS_T		
Rows (unadjusted)	$R-1 = 3$	SS_R	MS_R	
Columns (adjusted)	$C-1 = 2$	$SS_{C\text{-}A}$	$MS_{C\text{-}A}$	$F_1 = MS_{C\text{-}A}/MS_E$
Columns (unadjusted)	$C-1 = 2$	SS_C	MS_C	
Rows (adjusted)	$R-1 = 3$	$SS_{R\text{-}A}$	$MS_{R\text{-}A}$	$F_2 = MS_{R\text{-}A}/MS_E$
Interaction	$(R-1)(C-1) = 6$	SS_I	MS_I	$F_3 = MS_I/MS_E$
Error	$N-RC = 42$	SS_E	MS_E	

The first two F-ratios can be used to test the following hypotheses:

H_o: The "B" terms in the model are not needed; H_o: The "C" terms in the model are not needed. The third F-ratio is the same test for the interaction obtained in the preliminary AOV table. Notice that the SS for columns is obtained after correction for rows. That is, $SS_{C\text{-}A}$ (columns adjusted for rows) = SS_M (main effects in preliminary AOV table) - SS_{rows} (rows ignoring the column effects). Hence, some of the calculation for the final AOV by the method of fitting constants are derived from the preliminary AOV table.

In conclusion, the method of fitting constants allows us to make "good" tests for main effects if the interaction term is absent. Also, if one or more n_{ij} = zero we must use this method since the interpretation of a significant interaction is questionable anyway. After determining that the row and/or column means differ significantly, one might wish to do some type of multiple comparison procedure to determine where the significant differences lie.

Interpreting the Method of Squared Means AOV

When interaction is assumed present in our model or suspected to be present in the model after studying the preliminary AOV table, the method of squared means can be used to find "good" estimates of the main effects if all $n_{ij} > 0$. This analysis operates on the cell means weighted by $W_i = c^2/(\sum_j 1/n_{ij})$ for the ith row and $W_j = r^2/(\sum_i 1/n_{ij})$ for the jth column. The model for this situation would be:

$$Y = A + B * (\text{ROW LEVEL}) + C * (\text{COLUMN LEVEL}) +$$

$$D\,(\text{ROW, COLUMN LEVELS}) + \text{ERROR}$$

where A represents the average value and D represents the coefficient for the interaction term. The method, which is described on pages 24-29 of Bancroft, would yield an AOV table as follows:

Source	DF	SS	MS	F-Ratio
Total	$N-1 = 53$			
Rows (weighted)	$R-1 = 3$	$SS_{R\text{-}W}$	$MS_{R\text{-}W}$	$MS_{R\text{-}W}/MS_E$
Columns (weighted)	$C-1 = 2$	$SS_{C\text{-}W}$	$MS_{C\text{-}W}$	$MS_{C\text{-}W}/MS_E$
Interaction	$(R-1)(C-1) = 6$	SS_I	MS_I	MS_I/MS_E
Error	$N-RC = 42$	SS_E	MS_E	

The F-ratios for rows and columns using the weighted cell means will indicate if the main effects are significant. Of course, if the interaction term is already determined to be significant, the interpretation of the main effects must be given careful consideration. Quite frequently experimenters find it useful to plot the subclass means in order to study the "pattern" for the interaction.

Orthogonal Polynomial Breakdown

If the levels of the row and/or column factors are quantitative, it might be of interest to decompose the sum of squares for these terms into single-degree-of-freedom terms for a polynomial model. For example, suppose that the row levels are quantitative such as the temperature levels which we described above (100, 150, 175, 200°C). Since there are four levels, it is possible to fit up to a third degree polynomial to the row levels. Hence, the SS for rows could be decomposed into orthogonal components for linear, quadratic and cubic terms, each with one degree of freedom. The Orthogonal Polynomial special function key will perform the elaborate calculations necessary for this breakdown even if the row or column levels are unequally spaced. (For example, the column levels were given as 5, 10, 15 psi. Instead, they could have been 5, 10, 20 psi with unequal spacings between the levels.)

For further information about these procedures, see references 1, 2, and 3.

References

1. Bancroft, T.A. (1968). Topics in Intermediate Statistical Methods. The Iowa State University Press, Ames, Iowa.
2. Kopitzke, R. (and Boardman, T.J. Editor). 9830A Two-Way Unbalanced Analysis of Variance Pack. Hewlett-Packard, September, 1974, Part No. 09830-70852.
3. Searle, S.R. (1971). Linear Models, John Wiley and Sons.

Factorial Design

Object of Program

This program will calculate the complete analysis of variance table for a two-, three-, or four-factor, completely balanced experiment. There may be multiple observations per cell and the entire experiment may be replicated in blocks. The program will automatically print out all main effect and two-way interaction means. If three- or four-way interactions exist, these interaction means may be printed. If there is more than one observation per cell, then tests for homogeneity of variance may be computed. If the experiment has not been replicated, or only one observation per mean is present, there will be no F values computed. All F tests assume that the factors are fixed. A label of up to ten characters may be assigned to each factor.

Data may be entered two ways if there are multiple observations (samples) per cell. See the Data Structures section in this writeup for a complete description of the form of the data entry for Factorial Designs.

References

1. Cochran, W.G. and Cox, G.M., Experimental Designs, John Wiley and Sons, Inc.,1957.
2. Snedecor, G.W. and Cochran, W.G., Statistical Methods, Iowa State University Press, 1967.

User Instructions

1. The title will be printed.
2. When "Number of factors in design? (2, 3, or 4)" is displayed:
   a. Enter the number of factors. It must be 2, 3, or 4.
   b. Press: CONT
3. When "Number of levels of factor 'F'?" is displayed:
   a. Enter the number of levels of this factor.
   b. Press: CONT
   c. If more factors are in the design, go to step 3.
4. When "Number of blocks in this design?" is displayed:
   a. Enter the number of blocks.
   b. Press: CONT
5. When "Number of observations per treatment combination within each block?" is displayed:
   a. Enter the number of minor replications.
   b. Press: CONT

6. When "Is the above information correct?" is displayed:

   a. Press: NO if the design information is not correct.

   b. Go to step 2.

   or

   a. Press: YES if everything is correct.

7. When "Do you want to assign your own names to the factors?" is displayed:

   a. Press: NO if the default names are satisfactory.

   b. Go to step 9.

   or

   a. Press: YES if the factor names are to be specified.

8. When "Enter the name for factor 'F' (<11 characters)" is displayed:

   a. Enter the name to be associated with this factor.

   b. Press: CONT

   c. If more factors are to be named, go to step 8.

9. If subfiles are undefined, go to step 11.

10. When "Enter subfile to be used, or 0 if subfiles are to be ignored." is displayed:

    a. Enter the subfile number, or 0.

    b. Press: CONT

11. If there is only one variable in the data set, go to step 17.
12. If there is only one minor replication, go to step 16.
13. A description of the data entry modes will be displayed on the CRT.
14. When "Data entry option?" is displayed:

    a. Enter 1 if all the data is contained in one variable.

    b. Press: CONT

    c. Go to step 16.

    or

    a. Enter 2 if the minor replications are contained in more than one variable.

    b. Press: CONT

15. When "Enter the variable numbers of the minor replications, separated by commas?" is displayed:

    a. Enter the variable numbers that contain the minor replications. The number of variables must equal the number entered at step 5.

    b. Press: CONT

    c. Go to step 17.

16. When "Which variable number contains the response?" is displayed:

    a. Enter the variable number.

    b. Press: CONT

17. If the design specified is not compatible with the data supplied or the design is too large for the machine to handle, an error message will be displayed along with "SELECT KEY"

18. When "Enter the number of decimal places to be used when printing calculated values" is displayed:

    a. Enter any digit from 0 to 9 and that many decimal places will be reported. Any two-digit number will cause the output to be printed in exponential (floating point) format.

    b. Press: CONT

19. If a hard-copy printer is being used, a summary of the specified design will be printed.
20. The main effect means and all two-way interaction means will be printed.
21. If there are only two factors, go to step 25.
22. When "Should the 3-way means be printed?" is displayed:

    a. Press: YES and all the three-way interaction means will be printed.

    b. Go to step 23.

    or

    a. Press: NO and the means will not be printed.

23. If there are only three factors, go to step 25.
24. When "Should the 4-way means be printed?" is displayed:

    a. Press: YES and all the four-way interaction means will be printed.

    b. Go to step 25.

    or

    a. Press: NO and the means will not be printed.

25. The analysis of variance table will be printed. If the design has no major or minor replications, no F ratios will be computed. Otherwise, the F tests will assume all factors are fixed.
26. If there is only one minor replication, go to step 28.
27. When "Should test for homogeneity of variance be made?" is displayed:

    a. Press: YES and the results of the test will be printed.
    b. Go to step 28.

    or

    a. Press: NO and the test will not be made.

28. If there is only one variable defined in the original data set, go to step 31.
29. If no subfiles are specified: When "Do you want to specify a new variable for this design?" is displayed:

    a. Press: YES if the same analysis is to be performed on a new variable.
    b. Go to step 11.

    or

    a. Press: NO if no further responses are to be analyzed.
    b. Go to step 33.

30. If subfiles are specified: When "Do you want to specify a new variable for this design, subfile combination?" is displayed:

    a. Press: YES if the same analysis is to be performed on a different variable(s) within the same subfile.
    b. Go to step 11.

    or

    a. Press: NO if no further responses are to be analyzed in this subfile.

NOTE

A new subfile may be declared or a new key may be selected later.

---

31. If subfiles are undefined, go to step 33.
32. When "Do you want to specify a new subfile for this design?" is displayed:

    a. Press: YES if a new subfile and new variables are to be declared for this design.
    b. Go to step 10.

    or

    a. Press: NO if a new key is to be selected.

33. When "SELECT KEY" is displayed, you may press

a. Orthogonal Polynomials

b. Treatment Contrasts

c. Interaction Plots

or

d. Multiple Comparisons to do further analysis on this design, or, you may select another key.

Example

Twenty-four laboratory rats were deprived of food, except for one hour per day, for several weeks. At the end of that time, each rat was inoculated with one of four doses of a certain drug and, after one of three amounts of time, was fed. The weight (in grams) of the food ingested by each rat was measured. The purpose of the experiment is to determine the effect of the drug on the motivation of the rats.

A Time before feeding (hours)	**B** Dosage (mg/kg) .1	.3	.5	.7
1	9.077 8.77	5.63 8.76	4.42 3.01	1.38 3.96
5	9.16 11.82	11.57 11.53	5.22 9.21	5.72 4.69
9	16.08 14.65	10.37 14.46	7.27 6.10	5.48 9.28

The design for this experiment is a two-way factorial with three levels of time and four dosage levels of the drug. Two rats (observations) per experimental combination were used. The data can be subjected to an analysis of variance in order to determine if there are significant differences between the three times before feeding or the four dosages of the drug. In addition, we can determine if there is a significant interaction between time and dosage.

The F ratios indicate no significant interaction effect (F=.915), significant differences in time levels (F=14.819) and dosage levels (F=19.533). The orthogonal polynomial decomposition for the time factor (A) shows a significant linear effect. The decomposition for the dosage factor (B) shows a highly significant linear effect and a cubic effect.

Even though the AB interaction (time or dosage) is not significant, a plot of the two-way means was included to show results of the INTERACTION PLOT routine. A reference LSD value is shown on the interaction plot.

```
******************************************************************************
*                              DATA MANIPULATION                             *
******************************************************************************
                             FOOD DEPRIVATION OF RATS

Data file name:
Number of observations: 12
Number of variables: 2

Variables names:
   1. OBS 1 WT
   2. OBS 2 WT


                             FOOD DEPRIVATION OF RATS


     Variable # 1   Variable # 2
OBS#
   1
         9.07000        8.77000
   2
         5.63000        8.76000
   3
         4.42000        3.01000
   4
         1.38000        3.96000
   5
         9.16000       11.82000
   6
        11.57000       11.53000
   7
         5.22000        9.21000
   8
         5.72000        4.69000
   9
        16.08000       14.65000
  10
        10.37000       14.46000
  11
         7.27000        6.10000
  12
         5.48000        9.28000
```

```
***********************************************************************************
*                        FACTORIAL ANALYSIS OF VARIANCE                           *
***********************************************************************************
                              FOOD DEPRIVATION OF RATS


DESIGN
------
   Number of factors =  2
    No. of levels of factor A =  3
    No. of levels of factor B =  4
    No. of major replications (blocks) =  1
    No. of minor replications (samples) =  2

Subfiles will be ignored
Response variable(s) are :
Variable no.  1      OBS 1 WT
Variable no.  2      OBS 2 WT

MEANS
-----
* Overall mean =          8.23375


* Main Effect Means :

Factor A - TIME   Levels ( 1 - 3 ) :
         5.62500          8.61500          10.46125
Factor B - DOSAGE   Levels ( 1 - 4 ) :
        11.59167         10.38667           5.87167          5.08500


* Two Way Interaction Means :

Factor A - TIME  down and  Factor B - DOSAGE  across
                      1               2               3               4
    1               8.92000         7.19500         3.71500         2.67000
    2              10.49000        11.55000         7.21500         5.20500
    3              15.36500        12.41500         6.68500         7.38000



ANOVA TABLE
-----------
                 Factorial Analysis of Variance

Source (Name)     df    Sums of Squares        Mean Square   F Ratio  F-Prob
----------------------------------------------------------------------------
Total             23       339.96336            14.78102
A    TIME          2        95.30148            47.65074  14.819   .0006
B    DOSAGE        3       188.42831            62.80944  19.533   .0001
AB                 6        17.64783             2.94130    .915   .5168
Sampling Error    12        38.58575             3.21548
   NOTE: F tests assume that all factors are fixed
```

FACTOR LEVELS — CELL STATISTICS

Blk	A	B	Mean	Std Dev	Variance	Coef Var %
1	1	1	8.92000	.21213	.04500	2.38
1	1	2	7.19500	2.21324	4.89845	30.76
1	1	3	3.71500	.99702	.99405	26.84
1	1	4	2.67000	1.82434	3.32820	68.33
1	2	1	10.49000	1.88090	3.53780	17.93
1	2	2	11.55000	.02828	.00080	.24
1	2	3	7.21500	2.82136	7.96005	39.10
1	2	4	5.20500	.72832	.53045	13.99
1	3	1	15.36500	1.01116	1.02245	6.58
1	3	2	12.41500	2.89207	8.36405	23.29
1	3	3	6.68500	.82731	.68445	12.38
1	3	4	7.38000	2.68701	7.22000	36.41

Bartlett's test :
    Chi squared = 13.1835 with 11 degrees of freedom
    Prob(Chi squared >   13.1835) = .2820

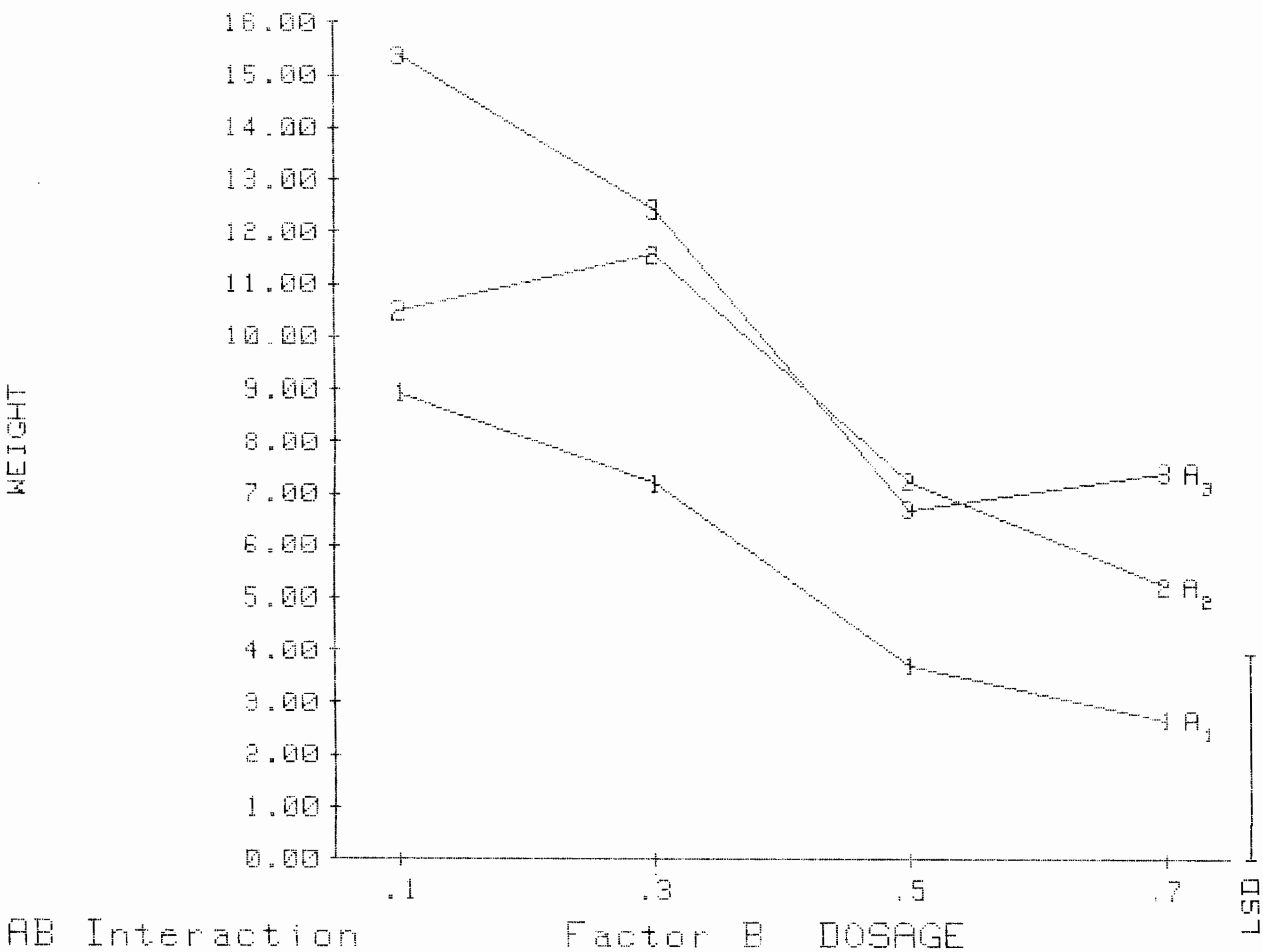
```
*******************************************************************************
*                            INTERACTION PLOT                                 *
*******************************************************************************
```

Last design was a 3X4 Factorial with 2 minor replicates.
Variable numbers that contain responses :
1 2

The AB interaction will be plotted.

Error Mean Square to be used is 3.21548
t = 2.179 LSD = 3.90733040255

FOOD DEPRIVATION OF RATS



Nested or Partially Nested Design

Object of Program

This program will calculate and print the AOV for any valid nested design. The program does this by computing a general factorial and then combining sums of squares to get the desired results. There can be up to five nested factors if samples are entered. This program does not allow the experiment to be replicated in blocks. The program will not compute any F ratios unless the design is a completely nested design. All non-nested main effects, main effect means, and two-way interactions will be printed. If there are any non-nested, three-way interaction means, they may be printed.

Data may be entered two ways if there are multiple observations per cell. See the Data Structures section for a complete description of the form of data entry for nested designs.

Limitations

No blocks are permitted in this design.

Possible Designs

All possible designs are displayed with arbitrary factors P, Q, R and S. In the program you will be asked to match your factors (A, B, etc.) with these arbitrary labels to obtain the design you desire. The following options are available.

Number of factors = 2
P
Q(P)
Number of factors = 3

1	2	3
P	P	P
Q(P)	Q	Q(P)
R(Q(P))	PQ	R
	R(PQ)	PR
		QR(P)

Number of factors = 4

1	2	3	4
P	P	P	P
Q(P)	Q	Q	Q(P)
R(Q(P))	R	PQ	R
S(R(Q(P)))	PQ	R(PQ)	PR
	PR	S	QR(P)
	QR	PS	S
	PQR	QS	PS
	S(PQR)	PQS	QS(P)
		RS(PQ)	RS
			PRS
			QRS(P)

References

1. C.R. Hicks "Fundamental Concepts in the Design of Experiments" 2nd edition. Holt, Rinehart and Winston, 1973.
2. D.C. Montgomery "Design and Analysis of Experiments". Wiley, 1976.

User Instructions

1. The title will be printed.
2. When "Number of factors in design? (2, 3 or 4)" is displayed:
   a. Enter the number of factors. It must be 2, 3, or 4 even if they are nested factors.
   b. Press: CONT
3. When "Number of levels of factor 'F'?" is displayed:
   a. Enter the number of levels for this factor. If the factor is to be nested, enter the number of levels it has among each category of the factor it is nested in.
   b. Press: CONT
   c. If more factors are in the design, go to step 3.
4. When "Number of samples?" is displayed:
   a. Enter the number of samples.
   b. Press: CONT
5. When "Is the above information correct?" is displayed:
   a. Press: NO if an error has been made.
   b. Go to step 2.

   or

   a. Press: YES if the displayed information is correct.
6. Using the arbitrary labels P, Q, R, & S, all possible designs of the size you have specified will be displayed.
7. If the number of factors is greater than 2:
   a. Enter the design number that corresponds to your analysis.
   b. Press: CONT

8. You will be asked to match your factors with the arbitrary labels (P, Q, R, S) in the design you specified. The questions will look something like this:

   "Which factor is (P,Q,R, or S): (possible choices)?"

This question will be repeated until sufficient information has been supplied.

9. When "Do you want to assign your own names to the factors?" is displayed:

   a. Press: NO if the default names are satisfactory.
   b. Go to step 11.

   or

   a. Press: YES if the factor names are to be specified.

10. When "Enter the name for factor 'F' (<11 characters)" is displayed:

    a. Enter the name to be associated with this factor.
    b. Press: CONT
    c. If more factors are in the design, go to step 10.

11. If subfiles are undefined, go to step 13.
12. When "Enter the subfile to be used, or 0 if subfiles are to be ignored." is displayed:

    a. Enter the subfile number, or 0.
    b. Press: CONT

13. If there is only one variable in the data set, go to step 19.
14. If there is only one sample observation, go to step 18.
15. A description of the data entry modes will be printed.
16. When "Data entry option?" is displayed:

    a. Enter 1 if all the data is contained in one variable.
    b. Press: CONT
    c. Go to step 18.

    or

    a. Enter 2 if the samples (minor replications) are spread across more than one variable.
    b. Press: CONT

17. When "Enter the variable numbers of the minor replications, separated by commas." is displayed:

    a. Enter the variable numbers that contain the sample observations. The number of variables must equal the number entered at step 4.

    b. Press: CONT

    c. Go to 19.

18. When "Which variable number contains the response?" is displayed:

    a. Enter the number corresponding to the response variable.

    b. Press: CONT

19. If the design specified is not compatible with your data or the design is too large for the program to handle, an error message will be displayed along with "SELECT KEY".

20. If the design specified and the observations specified are not compatible, an error message will be printed. Following the error message, select another special function key.

21. When "Enter the number of decimal places to be used when printing calculated values" is displayed:

    a. Enter any digit from 0 to 9 and that many decimal places will be reported. Any two digit numbers will cause the output to be printed in exponential (floating point) format.

    b. Press: CONT

22. If a hard-copy printer is specified, a summary of the design will be printed.

23. All main effect means and two-way interaction means will be printed.

NOTE

Some of the interaction means may not be useful in certain nested designs.

---

24. If there are less than three factors, go to step 26.

25. When "Should the 3-way means be printed?" is displayed:

    a. Press: YES and these means will be printed.

    b. Go to step 26

    or

    a. Press: NO if these means are not to be printed.

26. The analysis of variance table will be printed.

27. If there is only one variable defined in the original data set, go to step 30.

28. If subfiles exist:
When "Do you want to specify a new variable for this design, subfile combination?" is displayed:

a. Press: YES if the same analysis is to be done on a different variable(s) for the same subfile.

b. Go to step 13.

or

a. Press: NO if no further responses are to be analyzed in this subfile. Note that a new subfile may be declared or a new key may be selected later on.

b. Go to step 31.

29. If subfiles do not exist:
When "Do you want a new variable for this design?" is displayed:

a. Press: YES if the same analysis is to be done on a different variable or variables.

b. Go to step 13.

or

a. Press: NO if no further responses are to be analyzed for this design.

30. If subfiles are undefined, go to step 32.

31. When "Do you want to specify a new subfile for this design?" is displayed:

a. Press: YES if a new subfile and new variables are to be declared for this design.

b. Go to step 12.

or

a. Press: NO if a new key is to be selected.

32. When "SELECT KEY" is displayed you may press:

a. Orthogonal Polynomials

b. Treatment Contrasts

c. Interaction Plots

or

a. Multiple Comparisons
if you wish to do further analysis on this design.

Example

In order to compare two methods of display, a group of six new Thanksgiving greeting cards were selected. Eight stores were selected for the "promotional" display method and another eight stores were used for the "integrated" display method. For each of the two methods and each of eight stores per method, the same six card styles were compared using a response (Y) which measured dollar sales adjusted for store size. The data for each type of display, store, and greeting card style are shown below:

Display Method 1 - "Promotional" (A)

Stores (C)

Card Style (B)		1	2	3	4	5	6	7	8
	1	$1.21	1.49	1.76	1.52	0.65	1.96	1.21	1.57
	2	1.72	2.09	2.21	2.36	2.83	3.99	2.01	2.62
	3	1.72	1.44	1.84	0.91	1.30	7.61	2.01	3.27
	4	0.29	0.92	0.37	0.72	0.43	3.99	2.35	4.71
	5	1.44	2.09	1.84	2.36	1.96	3.26	2.01	1.70
	6	4.43	3.66	0.51	1.78	2.13	5.58	1.41	2.75

Display Method 2 - "Integrated" (A)

Stores (C)

Card Style (B)		9	10	11	12	13	14	15	16
	1	$2.60	2.21	1.44	1.20	1.21	3.03	2.79	1.18
	2	1.67	1.16	1.73	1.92	4.84	2.88	4.10	1.48
	3	3.67	0.78	1.46	1.65	3.23	1.92	4.51	1.48
	4	1.33	0.39	1.33	1.37	2.02	1.68	4.51	2.34
	5	3.33	1.16	1.86	1.92	3.23	2.64	3.96	2.22
	6	4.67	1.90	2.61	3.27	2.26	2.36	2.30	1.55

The mixed nested AOV for this model with factor A (display), factor C (stores) nested in factor A, and factor B (card style) crossed with A and C is shown below. The proper MS for testing differences between the two methods of display is C(A). Notice that the F ratio would be less than one = .42135/4.85529 indicating no significant difference between the methods as well as a considerable amount of store to store variation in the adjusted sales value. There does, however, appear to be significant differences between the population means for card types, i.e. F = 2.57257/.92726 = 2.77 which is significant at the .024 level.

A fairly standard procedure for the response variable Y considered here is to transform this response by $Y^* = \ln(Y+1)$ in order to achieve a more homogeneous and consistent response. The next analysis of variance is performed on this new response. The net result is that the F ratio for differences in card type means is even more highly significant (3.93 versus 2.77).

An LSD multiple comparison procedure was done on the six card styles. The results of this comparison show significant differences between style four and all others except style one with certain other differences existing as well. However, if one were looking for the highest adjusted daily sales, one should probably choose one of styles five, two, or six since they were not significantly different from one another but were different from the other styles (although three is questionably different).

THANKSGIVING GREETING CARD EVALUATION

VARIABLE # 1

I	OBS(I)	OBS(I+1)	OBS(I+2)	OBS(I+3)	OBS(I+4)
1	1.21000	1.49000	1.76000	1.52000	.65000
6	1.96000	1.21000	1.57000	1.72000	2.09000
11	2.21000	2.36000	2.83000	3.99000	2.01000
16	2.62000	1.72000	1.44000	1.84000	.91000
21	1.30000	7.61000	2.01000	3.27000	.29000
26	.92000	.37000	.72000	.43000	3.99000
31	2.35000	4.71000	1.44000	2.09000	1.84000
36	2.36000	1.96000	3.26000	2.01000	1.70000
41	4.43000	3.66000	.51000	1.78000	2.13000
46	5.58000	1.41000	2.75000	2.60000	2.21000
51	1.44000	1.20000	1.21000	3.03000	2.79000
56	1.18000	1.67000	1.16000	1.73000	1.92000
61	4.84000	2.88000	4.10000	1.48000	3.67000
66	.78000	1.46000	1.65000	3.23000	1.92000
71	4.51000	1.48000	1.33000	.39000	1.33000
76	1.37000	2.02000	1.68000	4.51000	2.34000
81	3.33000	1.16000	1.86000	1.92000	3.23000
86	2.64000	3.96000	2.22000	4.67000	1.90000
91	2.61000	3.27000	2.26000	2.36000	2.30000
96	1.55000				

* * * * * The data and related information are stored in AOVEX2:T14 * * * * *

```
********************************************************************************
*                        NESTED ANALYSIS OF VARIANCE                           *
********************************************************************************
```

THANKSGIVING GREETING CARD EVALUATION

```
DESIGN
------
 Number of factors =  3
    No. of levels of factor A =  2
    No. of levels of factor B =  6
    No. of levels of factor C =  8
    No. of minor replications (samples) =  1

Subfiles will be ignored
Response variable(s) are :
Variable no.  1       ADJ $SALE

MEANS
-----
* Overall mean =          2.23271
```

```
* Main Effect Means :

Factor A - DISPLAY   Levels ( 1 - 2 ) :
        2.16646          2.29896
Factor B - CARD STYLE   Levels ( 1 - 6 ) :
        1.68938          2.47563          2.42500          1.79688          2.31125
        2.69813
Factor C - STORES   Levels ( 1 - 8 ) :
        2.34000          1.60750          1.58000          1.74833          2.17417
        3.40833          2.76417          2.23917


* Two Way Interaction Means :

Factor A - DISPLAY  down and  Factor B - CARD STYLE  across
                  1              2              3              4
                  5              6
    1           1.42125        2.47875        2.51250        1.72250
                2.08250        2.78125
    2           1.95750        2.47250        2.33750        1.87125
                2.54000        2.61500

Factor A - DISPLAY  down and  Factor C - STORES  across
                  1              2              3              4
                  5              6              7              8
    1           1.80167        1.94833        1.42167        1.60833
                1.55000        4.39833        1.83333        2.77000
    2           2.87833        1.26667        1.73833        1.88833
                2.79833        2.41833        3.69500        1.70833

Factor B - CARD STYLE  down and  Factor C - STORES  across
                  1              2              3              4
                  5              6              7              8
    1           1.90500        1.85000        1.60000        1.36000
                 .93000        2.49500        2.00000        1.37500
    2           1.69500        1.62500        1.97000        2.14000
                3.83500        3.43500        3.05500        2.05000
    3           2.69500        1.11000        1.65000        1.28000
                2.26500        4.76500        3.26000        2.37500
    4            .81000         .65500         .85000        1.04500
                1.22500        2.83500        3.43000        3.52500
    5           2.38500        1.62500        1.85000        2.14000
                2.59500        2.95000        2.98500        1.96000
    6           4.55000        2.78000        1.56000        2.52500
                2.19500        3.97000        1.85500        2.15000
```

ANOVA TABLE

Nested Analysis of Variance

Source (Name)	df	Sums of Squares	Mean Square
Total	95	148.05410	1.55846
A DISPLAY	1	.42135	.42135
C(A)	14	67.97401	4.85529
B CARD STYLE	5	12.86285	2.57257
AB	5	1.88785	.37757
CB(A)	70	64.90804	.92726

```
The following transformation was performed: a*(X^b)+c
  where a =  1
        b =  1
        c =  1
        X is Variable # 1
        Transformed data is stored in Variable # 2 (LN(Y+1)).


The following transformation was performed: a*ln(bX)+c
  where a =  1
        b =  1
        c =  0
        X is Variable # 2
        Transformed data is stored in Variable # 2 (LN(Y+1)).
```

```
*********************************************************************************
*                          NESTED ANALYSIS OF VARIANCE                          *
*********************************************************************************
                       THANKSGIVING GREETING CARD EVALUATION
```

DESIGN

```
 Number of factors =  3
    No. of levels of factor A =  2
    No. of levels of factor B =  6
    No. of levels of factor C =  8
    No. of minor replications (samples) =  1

Subfiles will be ignored
Response variable(s) are :
Variable no.  2      LN(Y+1)
```

MEANS

```
* Overall mean =          1.10684




* Main Effect Means :

Factor A - DISPLAY   Levels ( 1 - 2 ) :
        1.07113          1.14256
Factor B - CARD STYLE   Levels ( 1 - 6 ) :
         .96211          1.20824          1.14025           .91050          1.17304
        1.24693
Factor C - STORES   Levels ( 1 - 8 ) :
        1.12364           .91076           .91441           .98173          1.08246
        1.42477          1.27979          1.13721


* Two Way Interaction Means :

Factor A - DISPLAY  down and  Factor B - CARD STYLE  across
                      1              2              3              4
                      5              6
    1               .87095        1.23072        1.14037         .83505
                   1.11320        1.23650
    2              1.05326        1.18577        1.14014         .98594
                   1.23288        1.25736

Factor A - DISPLAY  down and  Factor C - STORES  across
                      1              2              3              4
                      5              6              7              8
    1               .93881        1.04199         .83267         .92667
                    .87674        1.63105        1.03123        1.28987
    2              1.30847         .77952         .99614        1.03678
                   1.28819        1.21849        1.52834         .98455

Factor B - CARD STYLE  down and  Factor C - STORES  across
                      1              2              3              4
                      5              6              7              8
    1              1.03696        1.03928         .95361         .85636
                    .64688        1.23948        1.06268         .86162
    2               .99136         .94914        1.08529        1.14176
                   1.55380        1.48164        1.36559        1.09737
    3              1.27090         .73431         .97198         .81083
                   1.13756        1.61225        1.40425        1.17994
    4               .55026         .49081         .58034         .70261
                    .73147        1.29663        1.45776        1.47409
    5              1.17878         .94914        1.04731        1.14176
                   1.26370        1.37063        1.35167        1.08132
    6              1.71356        1.30186         .84791        1.23703
                   1.16138        1.54799        1.03677        1.12892
```

ANOVA TABLE

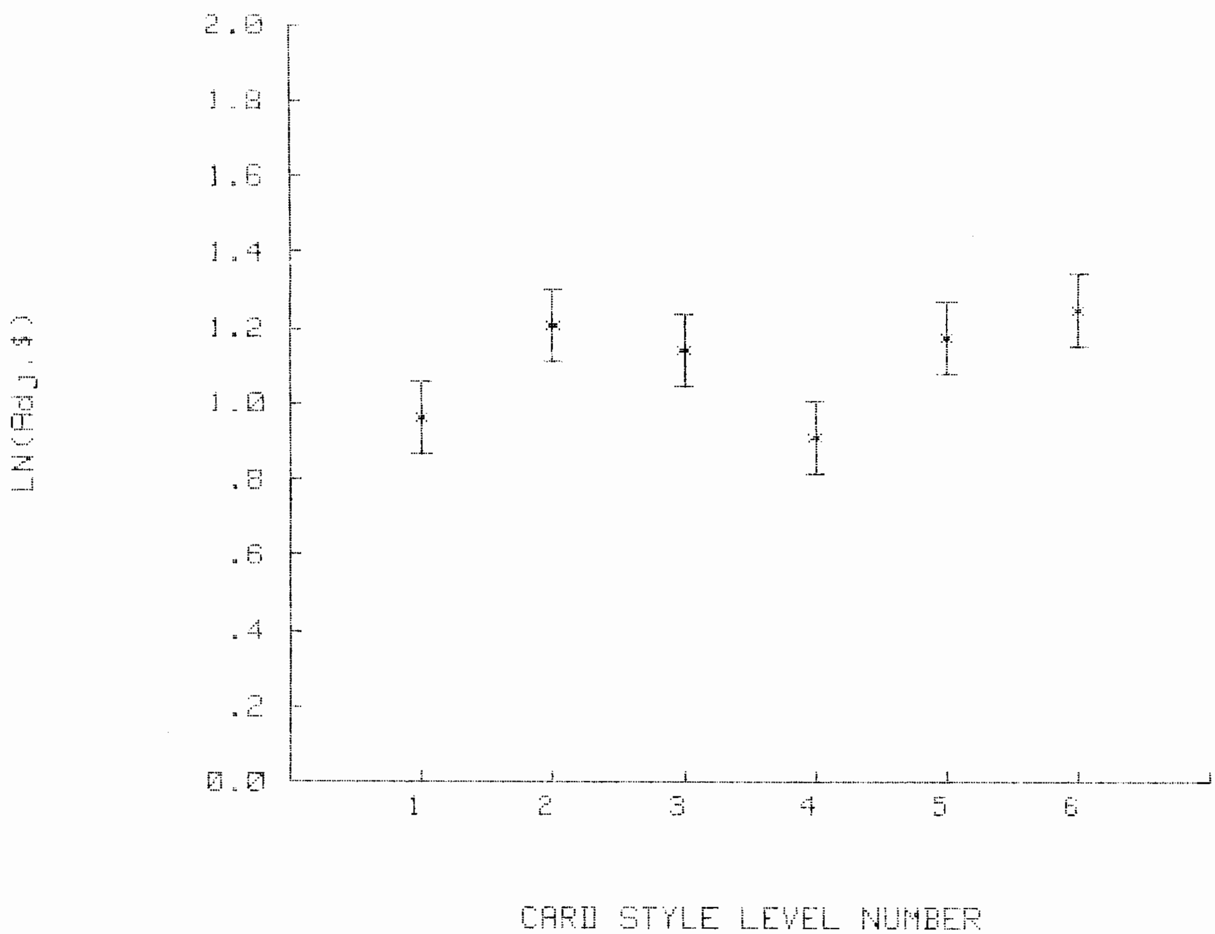
Nested Analysis of Variance

Source (Name)	df	Sums of Squares	Mean Square
Total	95	12.55307	.13214
A DISPLAY	1	.12245	.12245
C(A)	14	5.33726	.38123
B CARD STYLE	5	1.51847	.30369
AB	5	.16869	.03374
CB(A)	70	5.40619	.07723

Multiple Comparisons

MULTIPLE COMPARISON PLOT : LSD

THANKSGIVING GREETING CARD EVALUATION



Least Significant Difference

Error mean square = .07723
Degrees of freedom = 70
Harmonic average sample size = 16.0000
Alpha level = .05
Table value from Student's t = 1.99
LSD value = .195524462344

Multiple Comparisons on Factor CARD STYLE

Level	Mean	Sample Size	Separation
4	.91050	16	a
1	.96211	16	ab
3	1.14025	16	bc
5	1.17304	16	c
2	1.20824	16	c
6	1.24693	16	c

Split Plot Designs

Object of Program

This program will calculate a general factorial and then combine sums of squares to form specific error terms for the split plot or split-split plot design.

Blocks must be present and at least two factors are necessary. Up to three factors may be specified and minor replications (samples) may also be declared.

All main effects and interaction means will be printed. All computed F tests assume the factors are fixed.

Data may be entered two ways if there are multiple observations per cell. See the Data Structures section for a complete description of the form of data entry.

References

1. C.R. Hicks "Fundamental Concepts in the Design of Experiments" 2nd edition. Holt, Rinehart, Winston, 1973.
2. D.C. Montgomery "Design and Analysis of Experiments". Wiley, 1976.

User Instructions

1. The title will be printed.
2. When "Number of factors in design? (2 or 3)" is displayed:
   a. Enter the number of factors. It must be 2 or 3.
   Major factors that are in the split plots should be included.
   b. Press: CONT
3. When "Number of levels of factor 'F'?" is displayed:
   a. Enter the number of levels of this factor.
   b. Press: CONT
   c. If more factors are in the design, go to step 3.
4. When "Number of blocks in this design?" is displayed:
   a. Enter the number of blocks. There must be at least two.
   b. Press: CONT

5. When "Number of observations per treatment combination within each block?" is displayed:

   a. Enter the number of minor replications.

   b. Press: CONT

6. Any design that is too large for the program will cause an error message to be printed. Following the error message, select another key.

7. When "Do you want to assign your own names to the factors?" is displayed:

   a. Press: NO if the names are not to be changed.

   b. Go to step 9.

   or

   a. Press: YES if the factor names are to be specified.

8. When "Enter the name for factor 'F' (<11 characters)" is displayed:

   a. Enter the name to be associated with this factor.

   b. Press: CONT

   c. If more factors are in the design, go to step 8.

9. When "Which factor(s) are in the whole plots?" is displayed:

   a. Enter the letters of the factor names, separated by commas, that will be in the whole plot.

   b. Press: CONT

10. When "Which factor(s) are in the split-plots?" is displayed:

    a. Enter the letters of the factor name(s), separated by commas, that will be in the split plots.

    b. Press: CONT

11. When "Is the above information correct?" is displayed:

    a. Press: NO if the factor information is incorrect.

    b. Go to step 2.

    or

    a. Press: YES if the information is correct.

12. If the terms specified for the different plots effects are ambiguous, an error message will be printed and the program will return to step 9.

13. If subfiles are undefined, go to step 15.

14. When "Enter the subfile to be used, or 0 if subfiles are to be ignored." is displayed:

    a. Enter the subfile number, or zero.

    b. Press: CONT

15. If there is only one variable in the data set, go to step 21.
16. If there is only one minor replication, go to step 20.
17. A description of the data entry modes will be printed.
18. When "Data entry option?" is displayed:

    a. Enter 1 if all the data is contained in one variable.

    b. Press: CONT

    c. Go to step 20.

    or

    a. Enter 2 if the minor replications (samples) are spread across more than one variable.

    b. Press: CONT

19. When "Enter the variable numbers of the minor replications, separated by commas" is displayed:

    a. Enter the variable numbers that contain the sample observations. The number of variable numbers must equal the number entered at step 5.

    b. Press: CONT

    c. Go to step 21

20. When "Which variable number contains the response?" is displayed:

    a. Enter the number corresponding to the response variable.

    b. Press: CONT

21. If the design specified is not compatible with the data or the design is too large for the program, an error message will be displayed. Then, select another key.
22. When "Enter the number of decimal places to be used when printing calculated values." is displayed:

    a. Enter any digit from 0 to 9, and that many decimal places will be reported. Any two-digit number will cause the output to be printed in exponential (floating point) format.

    b. Press: CONT

23. If a hard-copy printer was specified, a summary of the design will be printed.
24. All the main effect means and two-way interaction means will be printed.

25. If there are only two factors, go to step 27.
26. When "Should the 3-way means be printed?" is displayed:

    a. Press: YES and the 3-way interaction means will be printed.
    b. Go to step 27.

    or

    a. Press: NO and the means will not be printed.

27. The analysis of variance table will be printed.
28. If the number of observations per cell is greater than one:
    When "Should test for homogeneity of variance be made?" is displayed:

    a. Press: YES if the test is desired.
    b. Go to step 29.

    or

    a. Press: NO if the test is not desired.

29. If there is only one variable defined in the original data set, go to step 32.
30. If subfiles do not exist:
    When "Do you want to specify a new variable for this design?" is displayed:

    a. Press: YES if the same analysis is to be performed on a different response.
    b. Go to step 15.

    or

    a. Press: NO if no further responses are to be analyzed in this design.
    b. Go to step 34.

31. If subfiles exist:
    When "Do you want to specify a new variable for this design, subfile combination?" is displayed:

    a. Press: YES if the same analysis is to be performed on a different response within the same subfile.
    b. Go to step 19.

    or

    a. Press: NO if no further responses are to be analyzed for this design, subfile combination.

NOTE

You will have a chance to specify a new subfile later.

---

32. If subfiles are undefined, go to step 34.

33. When "Do you want to specify a new subfile for this design?" is displayed:

    a. Press: YES if a new subfile is to be declared for this design.

    b. Go to step 14.

    or

    a. Press: NO if a new key is to be selected.

34. When "SELECT KEY" is displayed, you may press:

    a. Orthogonal Polynomials

    b. Treatment Contrasts

    c. Interaction Plots, or

    d. Multiple Comparisons
    to do further analysis on this design.

Example

Hicks (1973, ex. 13.1) describes a split-plot experiment in which four oven temperatures and three baking times were investigated with regard to the life, Y, of an electrical component. The oven temperatures and the replications (blocks) are in the whole plot while the baking times are in the subplots. Only one electrical component was subjected to the stress conditions within each block-baking time-temperature combination.

The data table is shown below:

		Oven Temp. B			
	Baking Time (A)	**580**	**600**	**620**	**640**
Replication 1	**5**	217	158	229	223
	10	233	138	186	227
	15	175	152	155	156
	5	188	126	160	201
2	**10**	201	130	170	181
	15	195	147	161	172
	5	162	122	167	182
3	**10**	110	185	181	201
	15	113	180	182	199

Since this is a balanced design with three replications, we need only use one variable for data entry. The data is entered across each row in the table above. Hence, three groups of replications are available with factor A as baking time and factor B as oven temperature.

Within the split-plot program, we answer that there are two factors and three major replications. The design is specified with factor B in the whole plot and factor A in the subplot. The F ratio shows only significant temperature effects (B). The HSD multiple comparison procedure suggests that oven temperature two is significantly lower in life time readings than are the other three temperatures.

This conclusion is supported, as should be expected, by the more 'liberal' LSD procedure shown on the next multiple comparison output.

If one runs this data set through the Factorial Analysis in order to separate the replication interaction terms as suggested by Hicks, one finds a highly questionable interaction between replications and baking time. To do this, you specify Factor A as replication, factor B as baking time, and factor C as oven temperature in the FACTORIAL program.

NOTE

On the printed AOV table, the mean square for AB (replication by baking time) is 1755.32 which is substantially larger than any of the other replication interactions.

After looking at the data set, we believe that Hicks may have rearranged the original data, since you would ordinarily not expect the replication interaction terms to differ by that much in a split plot. See if you agree.

```
*****************************************************************************
*                           DATA MANIPULATION                               *
*****************************************************************************
              HICKS(EX.13.1) SPLIT PLOT ON COMPONENT LIFE TIME

Data file name:
Number of observations: 36
Number of variables: 1

Variables names:
   1. LIFETIME


              HICKS(EX.13.1) SPLIT PLOT ON COMPONENT LIFE TIME


                               VARIABLE # 1
  I       OBS(I)        OBS(I+1)        OBS(I+2)        OBS(I+3)        OBS(I+4)
  1    217.00000       158.00000       229.00000       223.00000       233.00000
  6    138.00000       186.00000       227.00000       175.00000       152.00000
 11    155.00000       156.00000       180.00000       126.00000       160.00000
 16    201.00000       201.00000       130.00000       170.00000       181.00000
 21    195.00000       147.00000       161.00000       172.00000       162.00000
 26    122.00000       167.00000       182.00000       170.00000       185.00000
 31    181.00000       201.00000       213.00000       180.00000       182.00000
 36    199.00000
```

```
*****************************************************************************************
*                          SPLIT PLOT ANALYSIS OF VARIANCE                              *
*****************************************************************************************
                           HICKS(EX.13.1) SPLIT PLOT ON COMPONENT LIFE TIME


DESIGN
------
   Number of factors =  2
    No. of levels of factor A =  3
    No. of levels of factor B =  4
    No. of major replications (blocks) =  3
    No. of minor replications (samples) =  1

Subfiles will be ignored
Whole plot factor(s) are :
     Factor B
Split-plot factor(s) are :
     Factor A
Response variable(s) are :
Variable no.  1        LIFETIME

MEANS
-----
* Overall mean =           178.47


* Block and Main Effect Means :

Factor Blocks -        Levels ( 1 - 3 ) :
          187.42          169.33          178.67
Factor A - BakingTime     Levels ( 1 - 3 ) :
          177.92          183.58          173.92
Factor B - OVEN TEMP.     Levels ( 1 - 4 ) :
          194.89          148.67          176.78          193.56


* Two Way Interaction Means :

Factor A - BakingTime  down and  Factor B - OVEN TEMP.  across
                    1               2               3               4
    1            189.00          135.33          185.33          202.00
    2            201.33          151.00          179.00          203.00
    3            194.33          159.67          166.00          175.67
```

ANOVA TABLE

Split Plot Analysis of Variance

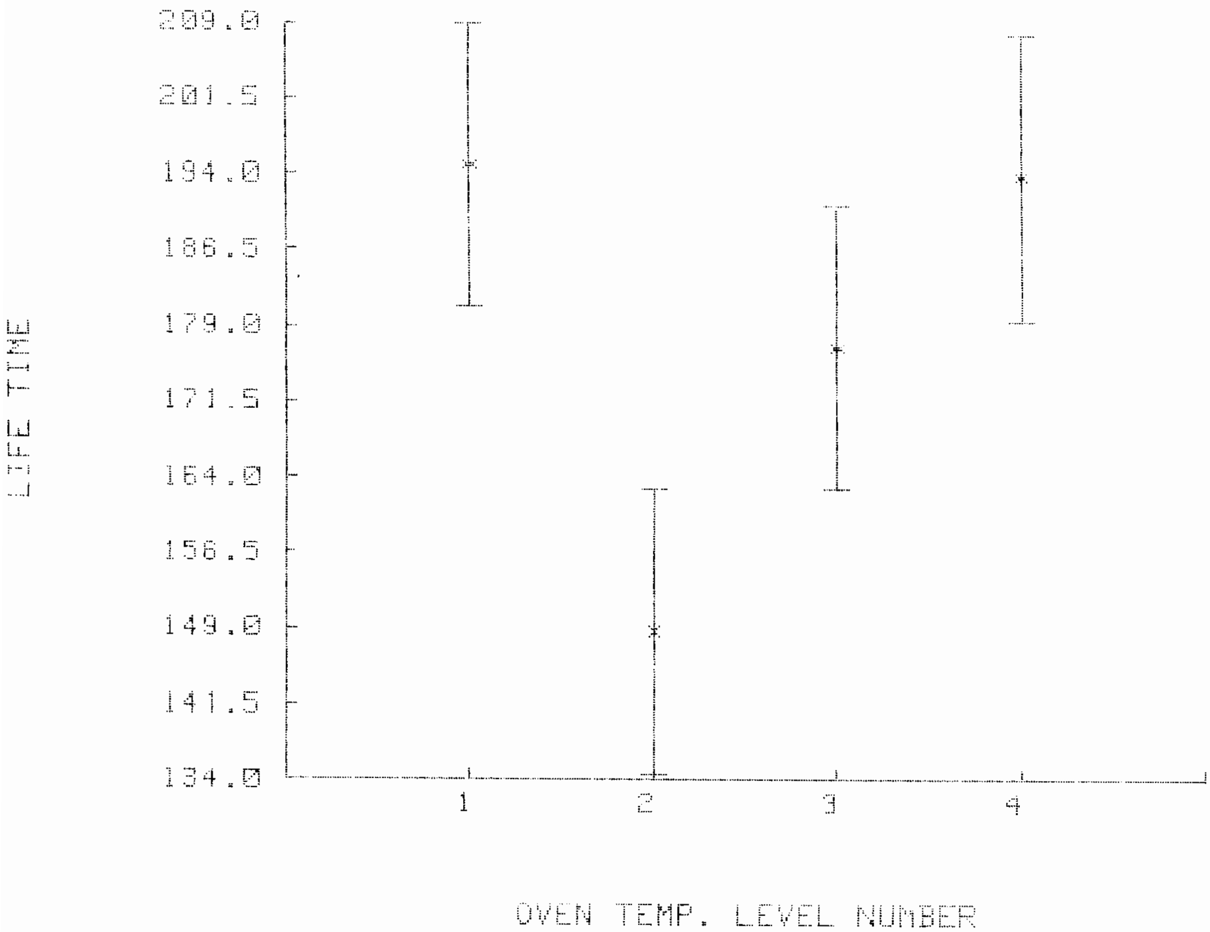
Source (Name)	df	Sums of Squares	Mean Square	F Ratio	F-Prob
Total	35	29330.97	838.03		
Blocks	2	1962.72	981.36	3.319	.1070
B OVEN TEMP.	3	12494.31	4164.77	14.086	.0040
Error (a)	6	1773.94	295.66		
A BakingTime	2	566.22	283.11	.456	.6418
BA	6	2600.44	433.41	.698	.6551
Error (b)	16	9933.33	620.83		

NOTE: F tests assume that all factors are fixed

****
Multiple Comparisons
****

MULTIPLE COMPARISON PLOT : TUKEY'S HSD

HICKS(EX.13.1) SPLIT PLOT ON COMPONENT LIFE TIME

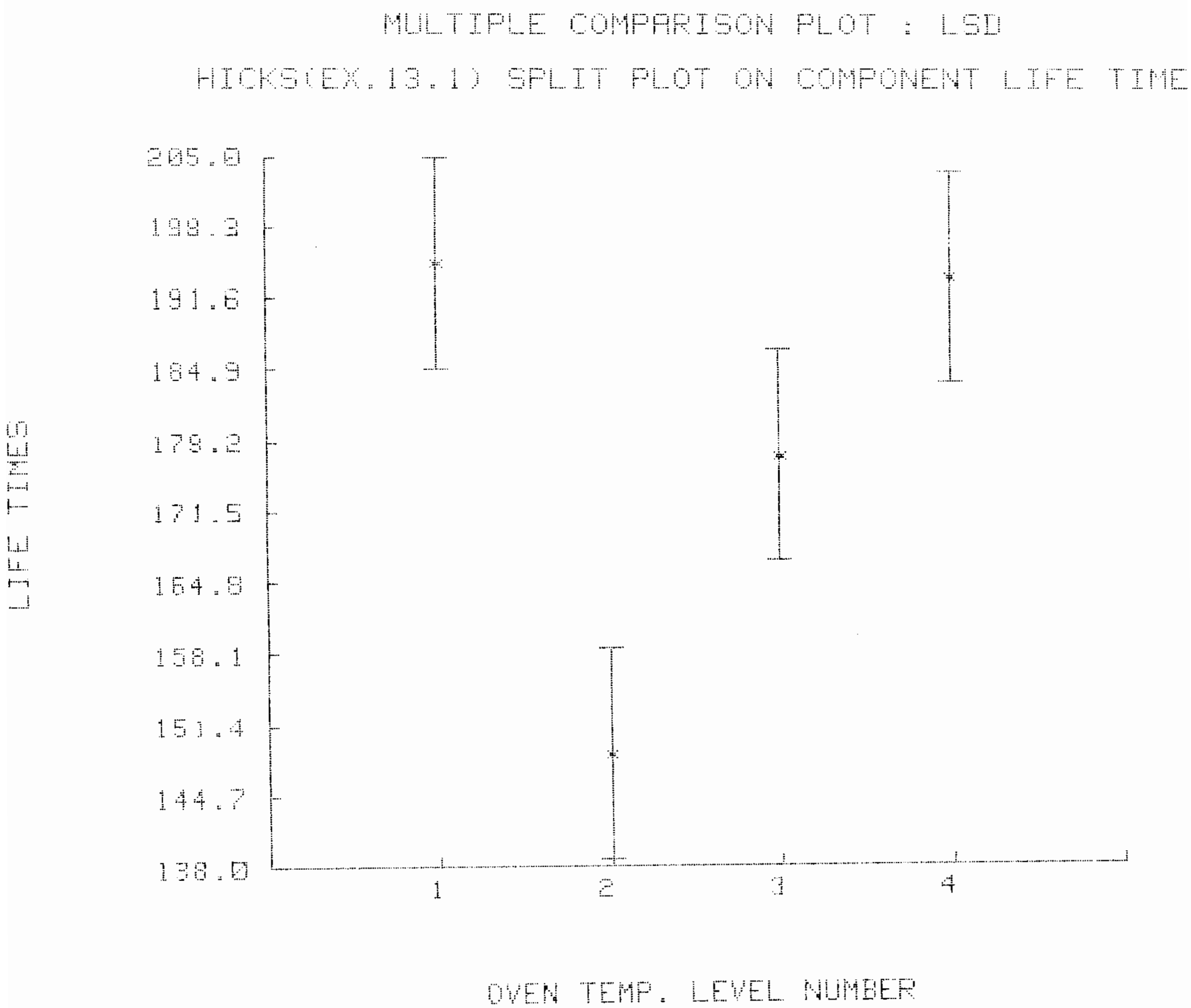


Tukey's HSD

Error mean square = 295.66
Degrees of freedom = 6
Harmonic average sample size = 9.0000
Alpha level = .05
Table value from Studentized range = 4.9
HSD value = 28.0847855212

Multiple Comparisons on Factor OVEN TEMP.

Level	Mean	Sample Size	Separation
2	148.67	9	a
3	176.78	9	b
4	193.56	9	b
1	194.89	9	b



Least Significant Difference

Error mean square = 295.66
Degrees of freedom = 6
Harmonic average sample size = 9.0000
Alpha level = .05
Table value from Student's t = 2.447
LSD value = 19.8346252181

Multiple Comparisons on Factor OVEN TEMP.

Level	Mean	Sample Size	Separation
2	148.67	9	a
3	176.78	9	b
4	193.56	9	b
1	194.89	9	b

```
*******************************************************************************
*                        FACTORIAL ANALYSIS OF VARIANCE                       *
*******************************************************************************
             HICKS(EX.13.1) SPLIT PLOT ON COMPONENT LIFE TIME
```

DESIGN

Number of factors = 3
No. of levels of factor A = 3
No. of levels of factor B = 3
No. of levels of factor C = 4
No. of major replications (blocks) = 1
No. of minor replications (samples) = 1

Subfiles will be ignored
Response variable(s) are :
Variable no. 1 LIFETIME

MEANS
-----

* Overall mean = 178.47

* Main Effect Means :

Factor A - REP Levels (1 - 3) :
187.42 169.33 178.67
Factor B - BAKE TIME Levels (1 - 3) :
177.92 183.58 173.92
Factor C - OVEN TEMP. Levels (1 - 4) :
194.89 148.67 176.78 193.56

* Two Way Interaction Means :

Factor A - REP down and Factor B - BAKE TIME across

	1	2	3
1	206.75	196.00	159.50
2	168.75	170.50	168.75
3	158.25	184.25	193.50

Factor A - REP down and Factor C - OVEN TEMP. across

	1	2	3	4
1	208.33	149.33	190.00	202.00
2	194.67	134.33	163.67	184.67
3	181.67	162.33	176.67	194.00

Factor B - BAKE TIME down and Factor C - OVEN TEMP. across

	1	2	3	4
1	189.00	135.33	185.33	202.00
2	201.33	151.00	179.00	203.00
3	194.33	159.67	166.00	175.67

* Three Way Interaction Means :

Factor A - REP, Level 1
Factor B - BAKE TIME down and Factor C - OVEN TEMP. across

	1	2	3	4
1	217.00	158.00	229.00	223.00
2	233.00	138.00	186.00	227.00
3	175.00	152.00	155.00	156.00

Factor A - REP, Level 2
Factor B - BAKE TIME down and Factor C - OVEN TEMP. across

	1	2	3	4
1	188.00	126.00	160.00	201.00
2	201.00	130.00	170.00	181.00
3	195.00	147.00	161.00	172.00

Factor A - REP, Level 3
Factor B - BAKE TIME down and Factor C - OVEN TEMP. across

	1	2	3	4
1	162.00	122.00	167.00	182.00
2	170.00	185.00	181.00	201.00
3	213.00	180.00	182.00	199.00

ANOVA TABLE

Factorial Analysis of Variance

Source (Name)	df	Sums of Squares	Mean Square
Total	35	29330.97	838.03
A REP	2	1962.72	981.36
B BAKE TIME	2	566.22	283.11
C OVEN TEMP.	3	12494.31	4164.77
AB	4	7021.28	1755.32
AC	6	1773.94	295.66
BC	6	2600.44	433.41
ABC	12	2912.06	242.67

One-Way Classification

Object of Program

This program will perform a one-way analysis of variance for treatments of equal or unequal size. You may give a ten character name to each treatment. For each treatment the name, sample size, total, mean, and standard deviation will be printed. The analysis of variance table will include all sums of squares and mean squares as well as the calculated F and the probability associated with getting that F value or one larger. You also have control over how many decimal places are to be printed on the output.

The treatments to be analyzed may be defined three different ways. See the Data Structures section of this manual for a detailed description of those methods.

References

1. W.J. Dixon, F.J. Massey "Introduction to Statistical Analysis" Third Edition. McGraw-Hill, 1969.
2. G.W. Snedecor, W.G. Cochran "Statistical Methods" Sixth Edition. Iowa State University Press, 1967.

User Instructions

1. The title will be printed.
2. When "How many treatments in this analysis?" is displayed:
   a. Enter the number of treatments to be analyzed. There must be two or more treatments.
   b. Press: CONT
3. When "Enter a name for the treatment/factor (<11 characters)" is displayed:
   a. Enter a name that the factor can be associated with.
   b. Press: CONT
4. A list of the default treatment names will be printed.
5. When "Do you want to assign your own names to the treatments?" is displayed:
   a. Press: NO if the default names are okay.
   b. Go to step 8.

   or

   a. Press: YES if the names of the treatments are to be changed.

6. When "Enter the name for treatment 'T' (<11 characters)" is displayed:

   a. Enter the name to be associated with this treatment.
   b. Press: CONT
   c. If more names are to be entered, go to step 6.

7. When "Are the names displayed on the CRT correct?" is displayed:

   a. Press: NO if changes are desired.
   b. Go to step 6.

   or

   a. Press: YES if no changes are necessary.

8. A description of the three methods for defining treatments will be printed.
9. When "Treatment definition mode?" is displayed:

   a. Enter 1 if variables are to be used as unique treatments.
   b. Press: CONT
   c. Go to step 12.

   or

   a. Enter 2 if the sample sizes for each treatment, beginning at either observation one or at the beginning of a subfile, are to be entered along with one variable number for the response.
   b. Press: CONT
   c. Go to step 11.

   or

   a. Enter 3 if contiguous groups of observations are to be used as treatments. The treatments are defined by entering the first and last observations of the given treatment.
   b. Press: CONT

10. When "Enter the first observation and the last observation number for treatment 'T' " is displayed:

    a. Enter the observation numbers in the original data set that correspond to this treatment.
    b. Press: CONT
    c. If more treatments are to be defined, go to step 10; otherwise, go to step 14.

11. When "Enter the number of observations in treatment 'T' " is displayed:

    a. Enter the number of observations in this treatment.
    b. Press: CONT
    c. If more treatments are to be defined, go to step 11; otherwise, go to step 12.

12. If subfiles are undefined, go to step 14.
13. When "Subfile # (enter 0 to ignore subfile) =?" is displayed:

    a. Enter the subfile number or zero.

    b. Press: CONT

14. Branch Point
    If treatment definition mode is:

    1 --go to step 15.

    2 --go to step 16.

    3 --go to step 16.

15. When "Enter the variable number to be used as treatment 'T' " is displayed:

    a. Enter the variable number to be used as this treatment.

    b. Press: CONT

    c. If more treatments are to be defined, go to step 15; otherwise, go to step 18.

16. If only one variable is defined in the data set, go to step 19.
17. When "Enter the variable number to be used as the response" is displayed:

    a. Enter the number corresponding to the response variable.

    b. Press: CONT

    c. Go to step 19.

18. When "Enter the name of the response" is displayed:

    a. Enter a name for the response.

    b. Press: CONT

19. When "Is the design description on the CRT correct?" is displayed:

    a. Press: YES if the information is correct.

    or

    a. Press: NO if changes are necessary.

    b. Go to step 9.

20. When "# of decimal places for printing calculated values (< = 7)" is displayed:

    a. Enter any digit from 0 to 7 and that many decimal places will be reported.

    b. Press: CONT

21. For each treatment, the treatment name, mean, standard deviation, and sample size will be printed.
22. The overall mean, standard deviation and number of observations will be printed.
23. The analysis of variance table will be printed.
24. A test for homogeneity of variance will be printed.
25. When "Do you wish to specify another response within this subfile?" is displayed:

    a. Press: YES if you would like to specify another response variable(s) for the treatment structure within this subfile.

    b. Go to step 14.

    or

    a. Press: NO otherwise.

NOTE

The underlined portion of the question in step 25 will not be displayed if mode 3 was selected or if subfiles were not defined.

---

26. If treatment definition Mode 3 was specified, go to step 28.
27. When "Do you wish to specify another subfile?" is displayed:

    a. Press: YES if you wish to change the subfile number.

    If treatment definition mode is:
    1--go to step 12.
    2--go to step 14.

    or

    a. Press: NO otherwise

28. When "SELECT KEY" is displayed, you may press:

    a. Orthogonal Polynomials

    b. Treatment Contrasts or

    c. Multiple Comparisons
    to do further analysis on these means.

Example

Tissue Culture Growth was studied after exposure to five 'sugar' treatments: control, 2% glucose, 2% fructose, 1% glucose and 1% fructose, and 2% sucrose. The response, Y, is the length (in ocular units) of pea section grown in tissue culture with auxin present.

The data was entered using One-Way AOV mode 2 in which all treatments are stored in one variable. Each treatment has ten observations (samples). Hence, observations 1 to 10 are in the first treatment, observations 11 to 20 are in the second treatment, etc. The F ratio for treatments shows a very strong indication that the population treatment levels are significantly different. Both the LSD and Duncan Multiple Comparison procedure separate the treatments into three non-overlapping groups - treatments 4, 3, and 2; and treatment 5; and treatment 1 (control). Hence, if you add either glucose (2) or fructose (3) or both (4) you get shorter lengths than if you use just sucrose which is in turn shorter than the control treatment.

```
*****************************************************************************************
*                              DATA MANIPULATION                                        *
*****************************************************************************************
                              TISSUE CULTURE GROWTH

Data file name:
Number of observations: 50
Number of variables: 1

Variables names:
   1. LENGTH
```

TISSUE CULTURE GROWTH

VARIABLE # 1

I	OBS(I)	OBS(I+1)	OBS(I+2)	OBS(I+3)	OBS(I+4)
1	75.00000	67.00000	70.00000	75.00000	65.00000
6	71.00000	67.00000	67.00000	76.00000	68.00000
11	57.00000	58.00000	60.00000	59.00000	62.00000
16	60.00000	60.00000	57.00000	59.00000	61.00000
21	58.00000	61.00000	56.00000	58.00000	57.00000
26	56.00000	61.00000	60.00000	57.00000	58.00000
31	58.00000	59.00000	58.00000	61.00000	57.00000
36	56.00000	58.00000	57.00000	57.00000	59.00000
41	62.00000	66.00000	65.00000	63.00000	64.00000
46	62.00000	65.00000	65.00000	62.00000	67.00000

* * * * * The data and related information are stored in AOVEX3:T14 * * * * *

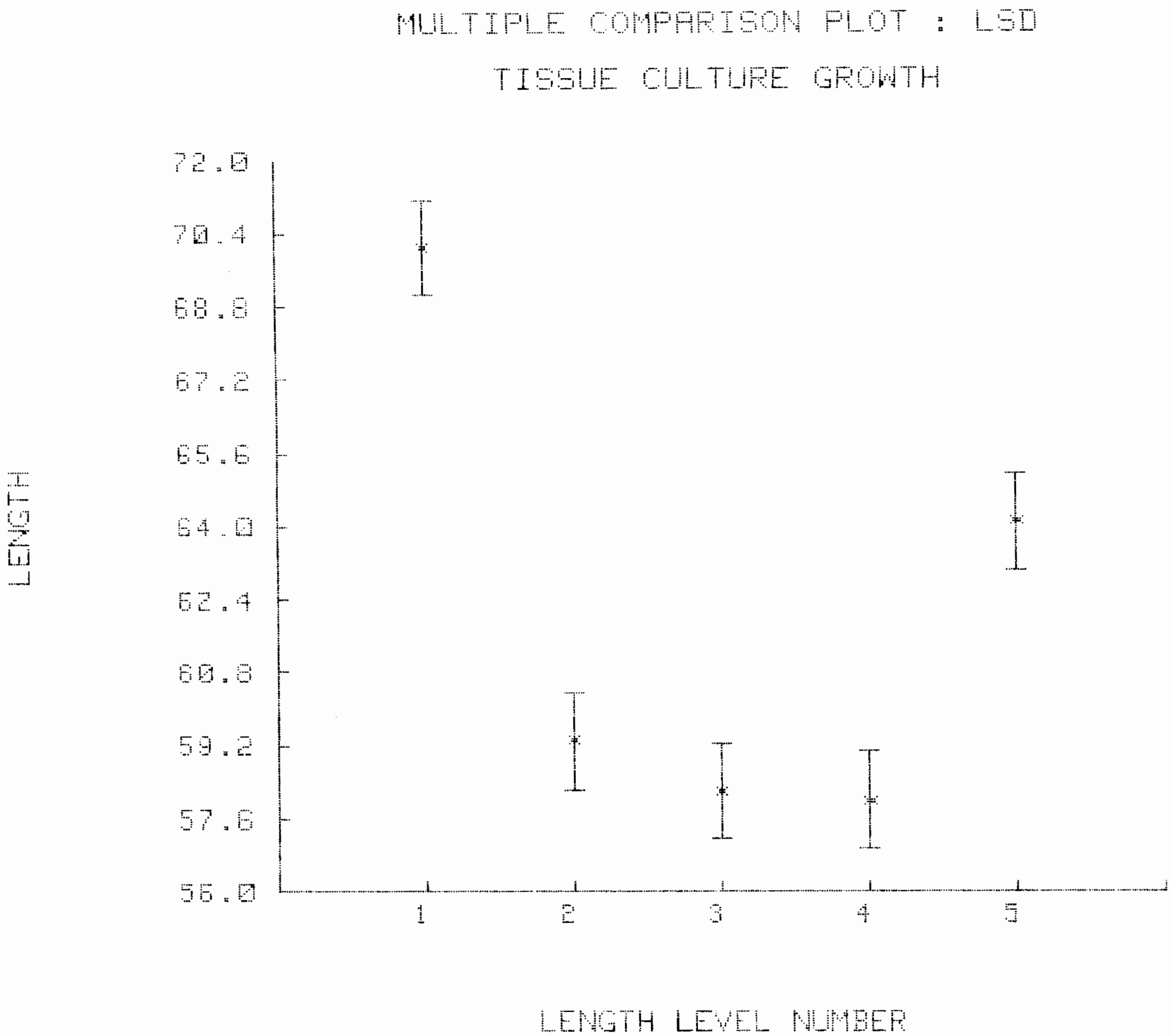
```
*******************************************************************************
                     ONE-WAY ANALYSIS OF VARIANCE:
                          TISSUE CULTURE GROWTH
*******************************************************************************
DESIGN
  # of treatments =  5
     # of observations in treatment 1  =  10
     # of observations in treatment 2  =  10
     # of observations in treatment 3  =  10
     # of observations in treatment 4  =  10
     # of observations in treatment 5  =  10
    Response = LENGTH


SUMMARY STATISTICS
                              Treatment Statistics
  Treatment name             Total                 Mean                Stan.Dev    N
   CONTROL              701.00000             70.10000              3.98469    10
   2% GLUCOSE           593.00000             59.30000              1.63639    10
   2% FRUCT.            582.00000             58.20000              1.87380    10
   1%GLU+FRU            580.00000             58.00000              1.41421    10
   2%SUCROSE            641.00000             64.10000              1.79196    10

   Overall             3097.00000             61.94000              5.19580    50
*******************************************************************************
ANOVA TABLE

                          One-Way Analysis of Variance Table
 Source          Df              SS                 MS          F-Ratio       F-Prob
 Total           49      1322.82000
 LENGTH           4      1077.32000          269.33000      49.36802       .00000
 Error           45       245.50000            5.45556
*******************************************************************************


Bartlett's test of homogeneity of variance : Chi-square value =  13.9386
with degrees of freedom =  4
```



Least Significant Difference

```
Error mean square = 5.45555555556
Degrees of freedom = 45
Harmonic average sample size =   10.0000
Alpha level = .05
Table value from Student's t = 2.014
LSD value = 2.10374915918
```

Multiple Comparisons on LENGTH

Level	Mean	Sample Size	Separation
4	58.00000	10	a
3	58.20000	10	a
2	59.30000	10	a
5	64.10000	10	b
1	70.10000	10	c

Duncan's Test

Error mean square = 5.45555555556
Degrees of freedom = 45
Alpha level = .05

Means Separated	Table Value	Required Difference
5	3.16000	2.33403
4	3.09500	2.28602
3	3.00500	2.21955
2	2.85000	2.10506

Multiple Comparisons on LENGTH

Level	Mean	Sample Size	Separation
4	58.00000	10	a
3	58.20000	10	a
2	59.30000	10	a
5	64.10000	10	b
1	70.10000	10	c

Two-Way Unbalanced Design

Object of Program

The purpose of this program is to perform an analysis of variance on a two-way classification with unequal subclass frequencies. The analysis may be performed in two ways.

If interactions are known to be present in the population, and all subclasses have at least one observation, then the method of weighted squares of means should be used to test the main effects.

If interactions are known to be absent in the population, or if at least one subclass has no observations, then the method of fitting constants should be used. In any case, if at least one subclass has no observations, the method of fitting constants must be used.

If it is not known whether or not interactions are present in the population, then a preliminary analysis of variance should be studied in order to test for interaction. If this test is significant, then the method of weighted squares of means should be used. A significance level of 0.25 may be used when testing for the presence of interaction.

User Instructions

1. The title is printed and the two forms of data storage are displayed.
2. When "Data storage type = " is displayed:

   a. Enter: "1" if three variables will be used to specify the response and factor levels (one for response, one for rows, one for columns).

   b. Press: CONT

   c. Go to step 3.

   or

   a. Enter: "2" if two variables will be used to specify the response and factor levels (one for response, one for packed identification).

   b. Press: CONT

   c. Go to step 4.

3. When "Variable numbers for row and column identifications (separated by a comma) = " is displayed:

   a. Enter the variable numbers corresponding to the row factor, and column factor with a comma in between.

   b. Press: CONT

   c. Go to step 5.

4. When "Variable number for packed identification = " is displayed:*

   a. Enter the variable number corresponding to the packed identification.

   b. Press: CONT

   c. Go to step 5.

***NOTE**

The form of the packing is as follows: You need four digits to specify the row and column numbers. The first two digits identify the level of the row factor, and the second two digits identify the level of the column factor. For example, 0203 stands for row level two and column level three. Do not use additional coding levels for your row or column identification. For example, if the row levels are really 10, 40, and 75, use 1, 2, and 3 to identify these levels.

---

5. When "Enter # of rows, # of columns (separated by a comma)" is displayed:

   a. Enter the number of rows and number of columns with a comma in between.

   b. Press: CONT

6. When "Do you wish to name the row factor?" is displayed:

   a. Press: NO if a name is not desired.

   b. Go to step 8.

   or

   a. Press: YES if you would like to name the row factor.

7. When "Enter the name of the row factor (<=10 characters)" is displayed:

   a. Enter the name of the row factor.

   b. Press: CONT

8. When "Do you wish to name the column factor?" is displayed:

   a. Press: NO if a name is not desired.

   b. Go to step 10.

   or

   a. Press: YES if you would like to name the column factor.

9. When "Enter the name of the column factor (<=10 characters)" is displayed:

   a. Enter the name of the column factor.

   b. Press: CONT

Example

The following data from Bancroft (1968, Ex. 1.3) is a two-way classification with factor A representing five different batches of silver and factor B representing two batches of iodine which are used to make silver iodide. The response, Y, is the reacting weights (coded). Apparently several samples were lost because the design is unbalanced.

		Iodine	
		I_1	I_2
Silver	S_1	22 25	−1 40 18
	S_2	41 41	23 13
	S_3	29 20 37	
	S_4	49 50	61
	S_5	55	

The data is entered using two variables. Variable one is used to identify the rows and columns and variable two contains the response, Y. Hence, a value in variable one of 0301 indicates that the observation in variable two is from the third level of silver (A) and the first level of Iodine (B). The Two-Way Unbalanced routine is used with the method of fitting constants selected as the desired procedures because of the presence of empty cells. This analysis indicates that the sampled batches of silver do not support the hypothesis of equality for the population means.

The multiple comparison procedure by Student, Newman & Keuls (SNK) shows no separation between the five samples of silver. This probably can be explained by both the conservative nature of the SNK procedure and the fact that the AOV procedure uses an adjusted mean square for silver.

```
*****************************************************************************
*                          DATA MANIPULATION                                *
*****************************************************************************
                CODED REACTING WEIGHTS OF SILVER AND IODINE

Data file name:
Number of observations: 16
Number of variables: 2

Variables names:
   1. ROW;COLUMN
   2. RWEIGHT
```

CODED REACTING WEIGHTS OF SILVER AND IODINE

OBS#	Variable # 1	Variable # 2
1	101.00000	22.00000
2	101.00000	25.00000
3	201.00000	41.00000
4	201.00000	41.00000
5	301.00000	29.00000
6	301.00000	20.00000
7	301.00000	37.00000
8	401.00000	49.00000
9	401.00000	50.00000
10	501.00000	55.00000
11	102.00000	-1.00000
12	102.00000	40.00000
13	102.00000	18.00000
14	202.00000	23.00000
15	202.00000	13.00000
16	402.00000	61.00000

* * * * * The data and related information are stored in AOVEX8:T14 * * * * *

****

TWO-WAY UNBALANCED ANALYSIS OF VARIANCE:
CODED REACTING WEIGHTS OF SILVER AND IODINE

****

DESIGN

# of rows = 5
# of columns = 2
Response = RWEIGHT

SUMMARY STATISTICS

Subclass Statistics

Row	Column	Total	Mean	Stan.Dev	N
1	1	47.0000	23.5000	2.1213	2
1	2	57.0000	19.0000	20.5183	3
2	1	82.0000	41.0000	0.0000	2
2	2	36.0000	18.0000	7.0711	2
3	1	86.0000	28.6667	8.5049	3
4	1	99.0000	49.5000	.7071	2
4	2	61.0000	61.0000	0.0000	1
5	1	55.0000	55.0000	0.0000	1

****

Row Statistics

Row	Total	Mean	N
1	104.0000	20.8000	5
2	118.0000	29.5000	4
3	86.0000	28.6667	3
4	160.0000	53.3333	3
5	55.0000	55.0000	1

****

Column Statistics

Col	Total	Mean	N
1	369.0000	36.9000	10
2	154.0000	25.6667	6

****

ANOVA TABLE

Preliminary AOV (Test two way model)

Source	Df	SS	MS	F-Ratio	F-Prob.
Total	15	4255.4375			
Subclass	7	3213.7708	459.1101	3.5260	.04908
Error	8	1041.6667	130.2083		

Preliminary AOV (Test for Interaction)

Source	Df	SS	MS	F-Ratio	F-Prob
Total	15	4255.4375			
Main	5	2722.2592	544.4518	4.1814	.03641
Int	2	491.5116	245.7558	1.8874	.21308
Error	8	1041.6667	130.2083		

Analysis of Variance (Method of Fitting Constants)

Source	Df	SS	MS	F-Ratio	F-Prob
Total	15	4255.4375			
SILVER	4	2572.3042	643.0760		
IODINE (Adj)	1	149.9550	149.9550	1.1517	.31450
IODINE	1	473.2042	473.2042		
SILVER (Adj)	4	2249.0550	562.2638	4.3182	.03749
Int	2	491.5116	245.7558		
Error	8	1041.6667	130.2083		

MULTIPLE COMPARISONS

Student-Newman-Keuls Test

Error mean square = 130.208333337
Degrees of freedom = 8
Alpha level = .05

Means Separated	Table Value	Required Difference
5	4.8900	36.3053
4	4.5300	33.6325
3	4.0400	29.9945
2	3.2600	24.2035

Multiple Comparisons on SILVER

Level	Mean	Sample Size	Separation
1	20.8000	5	a
3	28.6667	3	a
2	29.5000	4	a
4	53.3333	3	a
5	55.0000	1	a

10. When "Is the above information correct?" is displayed:

    a. Press: NO if changes are desired in the design specification.

    b. Go to step 2.

    or

    a. Press: YES if no changes are necessary.

10a. If there is only one subfile in the data set, go to step 12.

11. When "Subfile # (enter 0 to ignore subfile) =?" is displayed:

    a. Enter zero to ignore the subfiles.

    b. Press: CONT

    c. Go to step 12.

    or

    a. Enter the number of the subfile on which the two-way unbalanced AOV is to be performed.

    b. Press: CONT

12. When "Enter the variable number for the response" is displayed:

    a. Enter the variable number corresponding to the response.

    b. Press: CONT

13. When "Is the above information correct?" is displayed:

    a. Press: NO if changes are desired.

    b. Go to step 10a.

    or

    a. Press: YES if no changes are necessary.

14. When "# of decimal places for printing calculated values (<=7)?" is displayed:

    a. Enter the number of decimal places (0 to 7) desired for printing calculated values.

    b. Press: CONT

At this point, the program will print the total of all observations, the total number of observations and the overall mean. The cell statistics are printed next; the total, mean, number of observations, within cell variance and standard error of the mean are printed for each cell. Row and column statistics are then listed. The program also stores the means and number of observations in a scratch file for future use. Finally, the preliminary AOV is also printed.

15. When "Which method: Fitting constants (Enter 1); squared means (Enter 2)?" is displayed:

    a. Enter: "1" if the method of fitting constants is suggested.
    b. Press: CONT
    c. Go to step 16.

    or

    a. Enter: "2" if the method of squared means is suggested.
    b. Press: CONT

16. At this point, the program will print out the final AOV according to the method chosen in step 15.
17. If there is only one variable for a response variable in the data set, go to step 19.
18. When "Would you like to change response for this subfile?" is displayed:

    a. Press: YES if you would like to change the response for the same design.
    b. Go to step 12.

    or

    a. Press: NO if the same design on the new response is not desired.

19. If there are no subfiles in the data set, go to step 21.
20. When "Would you like to change subfiles?" is displayed:

    a. Press: YES if the same two-way AOV for a different subfile is desired.
    b. Go to step 11.

    or

    a. Press: NO if the same two-way AOV for a different subfile is not desired.

21. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys defined on the template.
    b. Go to the user instructions for the selected key.

One-Way Analysis of Covariance

Object of Program

This program will perform a one-way analysis of covariance for equal or unequal sample sizes. You may give a ten-character label to each treatment. For each treatment, a covariate (X) and a response variable (Y) must be specified.

For each treatment, the number of observations in the treatment, the means and standard deviations for the covariate (X) and the response (Y), the correlation between the two, and the equation of the least squares line will be printed. For the overall data, the same things will be computed and printed.

The corrected sums of squares tables will be printed and the analysis of covariance table with the calculated F and the probability associated with getting that F value or one larger will be printed.

Tests of the one-way analysis of variances for both X and Y, tests for equal slopes within treatments, and significant pooled regression will be calculated and printed.

The adjusted means and the standard errors of the adjusted means will be printed. These adjusted means will be saved for further analysis when doing multiple comparisons, or treatment contrasts.

Any time an observation is found with either the covariate (X) or response (Y) missing, the point will be deleted from the calculations.

You also have control over how many decimal places are to be printed on the output.

Treatments are defined in the same way as the one-way analysis of variance except that one variable number must contain the covariate (X) and another must have the response (Y). See The Data Structures section for further discussion.

References

1. W.J. Dixon, F.J. Massey "Introduction to Statistical Analysis", Third Edition. McGraw-Hill, 1969.
2. G.W. Snedecar, W.G. Cochran, "Statistical Methods", Sixth Edition. Iowa State University Press, 1967.

User Instructions

1. The title will be printed.
2. When "How many treatments in this analysis?" is displayed:

   a. Enter the number of treatments to be analyzed. There must be two or more treatments.

   b. Press: CONT

3. When “Enter a name for the treatment/factor (<11 characters)” is displayed:

   a. Enter a name for the factor.
   b. Press: CONT

4. A list of the default treatment names will be printed.
5. When “Do you want to assign your own names to the treatments?” is displayed:

   a. Press: NO if the default names are satisfactory.
   b. Go to step 8.

   or

   a. Press: YES if the names of the treatments are to be changed.

6. When “Enter the name for treatment ‘T’ (<=10 characters)” is displayed:

   a. Enter the name to be associated with this treatment.
   b. Press: CONT
   c. If more names are to be entered, go to step 6.

7. When “Are the names displayed on the CRT correct?” is displayed:

   a. Press: YES if no changes are necessary.
   b. Go to step 8.

   or

   a. Press: NO if changes are desired.
   b. Go to step 6.

8. A description of the three methods for defining treatments will be printed.
9. When “Treatment definition mode?” is displayed:

   a. Enter 1 if pairs of variables are to be used as unique treatments.
   b. Press: CONT
   c. Go to step 12.

   or

   a. Enter 2 if the sample sizes for each treatment, beginning at either observation one or at the beginning of a subfile are to be entered.
   b. Press: CONT
   c. Go to step 11.

   or

a. Enter 3 if contiguous groups of observations are to be used as treatments. These treatments are defined by entering the first and last observations of the given treatment.

b. Press: CONT

10. When "Enter the first observation and the last observation number for treatment 'T' " is displayed:

    a. Enter the observation numbers in the original data set that correspond to this treatment, separated by a comma.

    b. Press: CONT

    c. If more treatments are to be defined, go to step 10, otherwise go to step 14.

11. When "Enter the number of observations in treatment 'T' " is displayed:

    a. Enter the number of observations in this treatment.

    b. Press: CONT

    c. If more treatments are to be defined, go to 11, otherwise go to 12.

12. If subfiles are undefined, go to step 14.
13. When "Subfile # (enter 0 to ignore subfiles)=?" is displayed:

    a. Enter the subfile number or zero.

    b. Press: CONT

14. Branch Point
    If treatment definition mode is:
    1 -- Go to step 15.
    2 -- Go to step 16.
    3 -- Go to step 16.
15. When "Enter the X variable number, Y variable number for treatment 'T' " is displayed:

    a. Enter the variable numbers separated by a comma.

    b. Press: CONT

    c. If more treatments are to be defined, go to step 15; otherwise go to step 17.

16. When "Enter the X variable number, Y variable number used as the response" is displayed:

    a. Enter the variable numbers separated by a comma.

    b. Press: CONT

17. When "Is the design description on the CRT correct?" is displayed:

    a. Press: NO if changes are necessary.

    b. Go to step 9.

    or

    a. Press: YES if the design is correct.

18. When "# of decimal places to be used when printing calculated values" is displayed:

    a. Enter any digit from 0 to 7 and that many decimal places will be reported.

    b. Press: CONT

19. For each treatment, the name of the treatment, the number of observations in the treatment, the mean, and standard deviations for both the X and Y variables, the correlation coefficient and equation of the least squares line will be printed.

20. For all the data, the overall means and standard deviations for the X and Y variables, the number of observations in the experiment, the correlation coefficient and the equation of the least squares line will be printed.

21. The corrected sums of squares table will be printed.

22. F statistics with the corresponding probabilities will be printed for the one-way analysis of variance for both the X and Y variables, a test for whether the slopes of the regression lines within the groups are the same and a test for significant pooled regression will be calculated and printed.

23. The overall pooled slope with the standard deviation of the slope will be printed.

24. The analysis of covariance table will be printed.

25. The table of adjusted means for each treatment will be printed with the treatment name, the number of observations, the unadjusted Y mean, the adjusted mean and the standard error of the mean.

26. When "Do you wish to specify another response <u>within this subfile</u> ?" is displayed:

    a. Press: YES if you would like to specify another response for the treatment structure within this subfile, if any.

    b. Go to step 14.

    or

    a. Press: NO otherwise.

NOTE

The underlined portion of question 26 will not be displayed if mode 3 was selected or if subfiles are not defined.

---

27. If treatment definition mode 3 was chosen, go to step 30.

28. When "Do you wish to specify another subfile?" is displayed:

    a. Press: YES if you wish to change the subfile number.

29. Branch Point
    If treatment definition mode is:
    1 -- Go to step 12.
    2 -- Go to step 19.

    b. Press: NO otherwise.

30. When "SELECT KEY" is displayed, you may press:

    a. Orthogonal Polynomials

    b. Treatment Contrasts, or

    c. Multiple Comparison
    to do further analysis on these means.

Analysis of Covariance Example

An experiment to evaluate the effects of various growth stimulants on tomato seedlings was performed in which:

X = Initial length of seedling (m.m.)
Y = Growth in length (m.m.) during experiment

Stimulant X-4		Stimulant BC		Stimulant F32		Stimulant OX	
X	Y	X	Y	X	Y	X	Y
29	22	15	30	16	12	5	23
20	22	9	32	31	8	25	31
14	20	1	26	26	13	16	28
21	24	6	25	35	25	10	26
6	12	19	37	12	7	24	33

The data was entered using the first mode of storage for the covariance program. That is, each X,Y pair was stored in two variables and each of the four treatments used different variable pairs. Hence, for the Stimulant X-4, the initial length, X, was stored in Variable 1 and the growth, Y, was stored in Variable 2; while for the stimulant OX, the X value was stored in Variable 7 and the Y in Variable 8. Each variable has five observations.

The first part of the output from the One-way Covariance routines shows the within treatment statistics including totals, means, standard deviations, sample sizes, correlation coefficients, and regression coefficients. Note that the correlation coefficient and regression coefficient are for all of the data points taken together without regard to treatment group. Hence, it should not be surprising that no overall relationship exists between the X and Y variables. The test for homogeneity of regression coefficients confirms that we can accept the hypothesis that all treatment regression coefficients are essentially the same. The test for significance of pooled regression confirms that the relationship between the X and Y pooled across all treatments is significant (level = .0003).

Whereas the F ratio for treatment differences on the X's is non-significant (level = .12117), the F ratio on the original Y's is significant at the .00037 level. The analysis of covariance adjustment to the original data does not change the significance of the treatment effect (α = .00000), but rather makes the differences in the means even more pronounced. This is shown by studying the "Table of Means" and noting the adjustment made in the original Y means after the use of the covariate X.

The use of the Tukey HSD multiple comparison procedure shows that stimulants one and three differ from all other stimulants, while no significant difference can be shown between two and four.

```
********************************************************************************
*                            DATA MANIPULATION                                 *
********************************************************************************
           EFFECTS OF GROWTH STIMULANTS ON TOMATO SEEDLING LENGTHS

Data file name:
Number of observations: 5
Number of variables: 8

Variables names:
   1. X-4:I
   2. X-4:G
   3. BC:I
   4. BC:G
   5. F32:I
   6. F32:G
   7. OX:I
   8. OX:G
```

EFFECTS OF GROWTH STIMULANTS ON TOMATO SEEDLING LENGTHS

OBS#	Variable # 1	Variable # 2	Variable # 3	Variable # 4	Variable # 5	Variable # 6	Variable # 7	Variable # 8
1	29.00000	22.00000	15.00000	30.00000	16.00000	12.00000	5.00000	23.00000
2	20.00000	22.00000	9.00000	32.00000	31.00000	8.00000	25.00000	31.00000
3	14.00000	20.00000	1.00000	26.00000	26.00000	13.00000	16.00000	28.00000
4	21.00000	24.00000	6.00000	25.00000	35.00000	25.00000	10.00000	26.00000
5	6.00000	12.00000	19.00000	37.00000	12.00000	7.00000	24.00000	33.00000

* * * * * The data and related information are stored in AOVEX4:T14 * * * * *

****

ONE-WAY ANALYSIS OF COVARIANCE

EFFECTS OF GROWTH STIMULANTS ON TOMATO SEEDLING LENGTHS

****

DESIGN

# of treatments = 4
- # of observations in treatment 1 = 5
- # of observations in treatment 2 = 5
- # of observations in treatment 3 = 5
- # of observations in treatment 4 = 5

Covariate X = X-4:I

Response Y = X-4:G

SUMMARY STATISTICS

Treatment Statistics

Treatment		Total	Mean	Stan.Dev	N
X-4	X	90.0000	18.0000	8.5732	5
	Y	100.0000	20.0000	4.6904	5
BC	X	50.0000	10.0000	7.1414	5
	Y	150.0000	30.0000	4.8477	5
F32	X	120.0000	24.0000	9.7724	5
	Y	65.0000	13.0000	7.1764	5
OX	X	80.0000	16.0000	8.6891	5
	Y	141.0000	28.2000	3.9623	5

Overall	X	340.0000	17.0000	9.4088	20
	Y	456.0000	22.8000	8.5076	20

Within Treatment Regressions

Treatment	Corr.Coef.	Regression Coef.
X-4	.8331	.4558
BC	.8449	.5735
F32	.6310	.4634
OX	.9730	.4437
Total	-.0487	-.0440

****

ANOVA TABLE

One-Way Analysis of Variance Table(X-Variable)

Source	Df	SS	MS	F-Ratio	F-Prob
Total	19	1682.0000			
Treatment	3	500.0000	166.6667	2.2561	.12117
Error	16	1182.0000	73.8750		

One-Way Analysis of Variance Table(Y-Variable)

Source	Df	SS	MS	F-Ratio	F-Prob
Total	19	1375.2000			
Treatment	3	924.4000	308.1333	10.9364	.00037
Error	16	450.8000	28.1750		

****

Test of homogeneity of regression coefficients :
F-value = .0538 with 3 and 12 degrees of freedom
P(F> .0538) = .9828

Test of significance of pooled regression coefficient :
F-value = 21.8324 with 1 and 15 degrees of freedom
P(F> 21.8324) = .0003

Pooled Regression Coefficient = .4755
Pooled Correlation Coefficient = .7699

****

One Way Analysis of Covariance Table

Source	Df	SS	MS	F-Ratio	F-Prob
Total	18	1371.9444			
Treatment	3	1188.3559	396.1186	32.3647	.00000
Error	15	183.5885	12.2392		

****

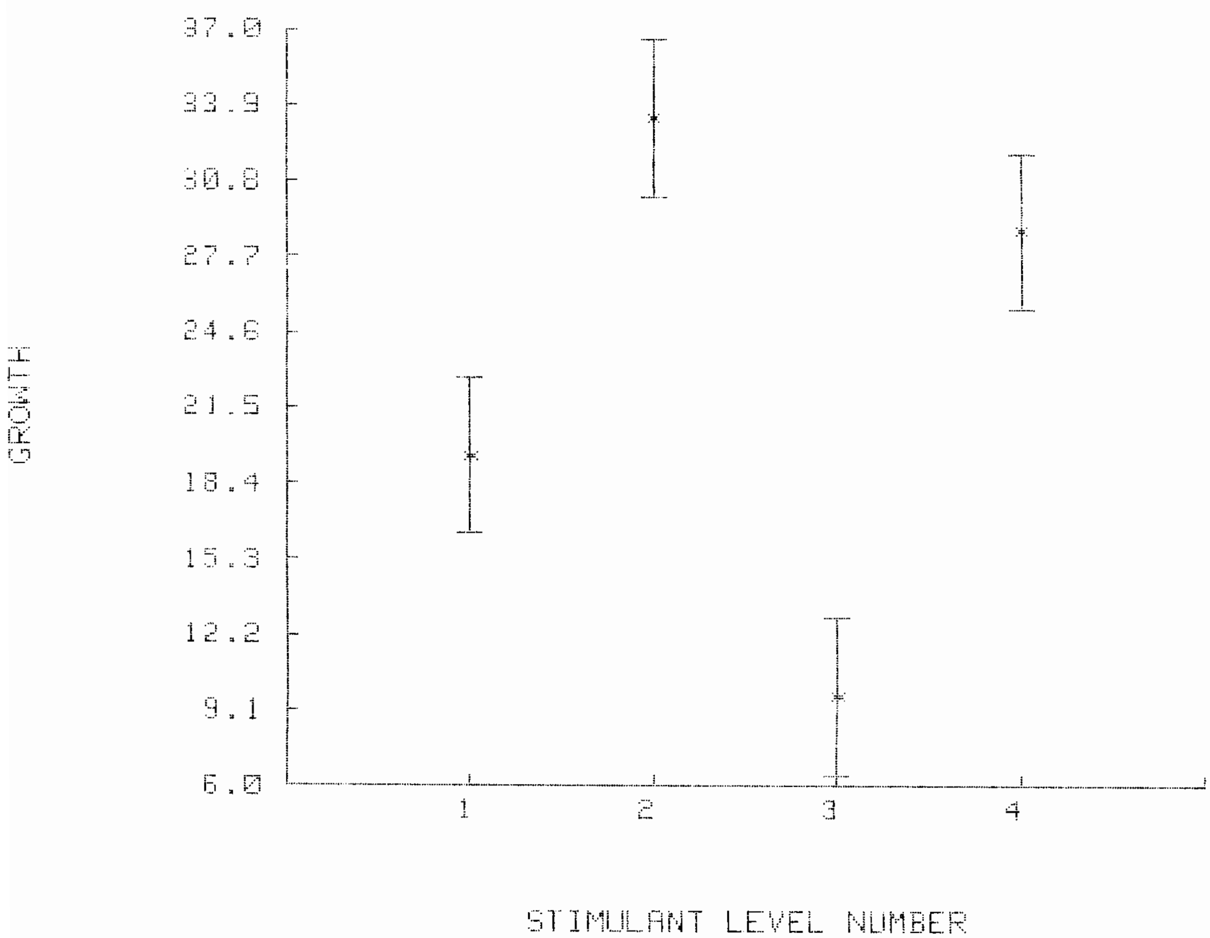
Table of Y Means

Treatment name	Unadjusted Y Mean	Adjusted Y Mean	Stand. Dev	N
X-4	20.0000	19.5245	1.5646	5
BC	30.0000	33.3283	1.5646	5
F32	13.0000	9.6717	1.5646	5
OX	28.2000	28.6755	1.5646	5

****
* MULTIPLE COMPARISONS *
****

MULTIPLE COMPARISON PLOT : TUKEY'S HSD

EFFECTS OF GROWTH STIMULANTS ON TOMATO SEEDLING LENGTHS



Tukey's HSD

```
Error mean square = 12.2392329385
Degrees of freedom = 15
Harmonic average sample size =     5.0000
Alpha level = .05
Table value from Studentized range = 4.08
HSD value = 6.38340296698
```

Multiple Comparisons on STIMULANT

Level	Mean	Sample Size	Separation
3	9.6717	5	a
1	19.5245	5	b
4	28.6755	5	c
2	33.3283	5	c

F-Prob

Object of Program

Given the numerator degrees of freedom, and the denominator degrees of freedom, and an F value>1, this program will calculate the probability that an F random variable has a value greater than or equal to the given F value.

References

1. Boardman, T.J. (editor) 9830A Statistical Distribution Pac, Hewlett-Packard (PN 09830-70854), September, 1974.
2. Boardman, T.J. (editor) 9845A General Statistics Package.
3. Boardman, T.J. and R.W. Kopitzke, "Probability and Table Values for Statistical Distributions", 1975, Proceedings of the Statistical Computing Section of The American Statistical Association, pp 81-86.

User Instructions

1. Press: The FPROB key.
2. When "Enter d.f. numerator, d.f. denominator, F value?" is displayed:
   a. Enter the three values separated by commas.
   b. Press: CONT
3. When "Another FPROB?" is displayed:
   a. Press: YES if another F probability is desired.
   b. Go to step 2.

   or

   a. Press: NO otherwise.
4. When "SELECT KEY" is displayed:
   a. Choose any special function key and go to the instructions for the selected key.

Orthogonal Polynomial

Object of Program

This program generates orthogonal polynomials. This allows you to determine if quantitative factor levels with equal or unequal spacings in the levels are linear, quadratic, etc., in their relationship to the response variable. The output includes the sum of squares, the F-ratio and the P(F>comp F) for each degree polynomial.

User Instructions

1. When "Is the design displayed on the CRT the latest one?" is displayed:

   a. Press: NO if the orthogonal polynomial will not perform on the design displayed on the CRT.

   b. Go to step 19.

   or

   a. Press: YES if the design displayed on the CRT is the latest one.

2. For a one-way classification design, go to step 3.
   For a two-way unbalanced classification design, go to step 6.
   For other designs, go to step 12.

3. When "Enter the max degree of orthogonal poly (<=K)" is displayed:

   a. Enter the maximum degree of orthogonal polynomial desired (it must be less than the number of treatments).

   b. Press: CONT

   or

   a. Enter zero if the treatments are not a quantitative variable or if it is not desired to do orthogonal polynomial comparisons on the treatments.

   b. Press: CONT

   c. Go to step 19.

4. When "Enter the level associated with treatment # 'i' " is displayed:

   a. Enter the quantity corresponding to treatment i.

   b. Press: CONT

   c. Repeat step 4 until all the levels have been entered.

5. When "Is the above information correct?" is displayed:

   a. Press: NO if changes are desired.

   b. Go to step 4.

or

a. Press: YES if no changes are necessary.

b. The program will print the orthogonal coefficients and the decomposition of the AOV table into linear effects, quadratic effects, etc.

c. Go to step 19.

6. When "Enter the max degree of row (factor A) orthogonal poly (<=K)" is displayed:

a. Enter the maximum degree desired for row orthogonal polynomial comparisons (it must be less than the number of levels of A).

b. Press: CONT

c. Go to step 7.

or

a. Enter zero if the rows are not a quantitative variable or if it is not desired to do orthogonal polynomial comparisons on the rows.

b. Press: CONT

c. Go to step 9.

7. When "Enter the level associated with row # 'T' " is displayed:

a. Enter the quantity corresponding to row T.

b. Press: CONT

c. Repeat step 7 until all the row levels have been entered.

8. When "Is the above information correct?" is displayed:

a. Press: NO if changes are desired.

b. Go to step 7.

or

a. Press: YES if no changes are necessary.

9. When "Enter the max. degree of column (factor B) orthogonal poly (<=K)" is displayed:

a. Enter the maximum degree desired for column orthogonal polynomial comparisons (it must be less than the number of levels of B).

b. Press: CONT

c. Go to 10.

or

a. Enter zero if the column is not a quantitative variable or if it is not desired to do polynomial comparisons on the columns.

b. Press: CONT

c. Go to step 19.

10. When "Enter the level associated with column # 'I' " is displayed:

a. Enter the quantity corresponding to column I.

b. Press: CONT

c. Repeat step 10 until all the column levels have been entered.

11. When "Is the above information correct?" is displayed:

a. Press: NO if changes are desired.

b. Go to step 10.

or

a. Press: YES if no changes are necessary.

b. Orthogonal coefficients and the decomposition of AOV table will be printed.

c. Go to step 19.

12. When "Are orthogonal poly comparisons on factor (K) desired?" is displayed:

a. Press NO if the orthogonal polynomial comparisons on factor (K) are not desired.

b. Go to step 12.

or

a. Press: YES if the orthogonal polynomial comparisons on factor (K) are desired.

13. When "Enter the max. degree of orthogonal poly (<=K)" is displayed:

a. Enter the maximum degree of othogonal polynomial comparison desired (it must be less than the number of levels of factor (K)).

b. Press: CONT

c. Go to step 12.

or

a. Enter zero if factor (K) is not a quantitative variable or if it is not desired to do orthogonal polynomial comparisons on factor (K).

b. Press: CONT

c. Go to step 18.

14. When "Enter the quantity associated with level # 'i' of factor 'K' " is displayed:

    a. Enter the quantity corresponding to level i of factor K.
    b. Press: CONT
    c. Repeat step 14 until all the levels of factor K have been entered.

15. When "Is the above information correct?" is displayed:

    a. Press: NO if changes are desired.
    b. Go to step 14.

    or

    a. Press: YES if no changes are necessary.

16. For designs other than nested or split-plot, go to step 18.
17. When "Enter error mean square and associated degrees of freedom" is displayed:

    a. Enter the error mean square and associated degrees of freedom to be used in this orthogonal polynomial analysis.
    b. Press: CONT

18. When "More orthogonal poly comparisons on another factor (YES/NO)?" is displayed:

    a. Press: YES if orthogonal polynomial comparisons on another factor are desired.
    b. Go to step 12.

    or

    a. Press: NO if orthogonal polynomial comparisons on another factor are not desired.

19. When "SELECT ANY KEY" is displayed:

    a. Press: Any of the keys defined on the template.
    b. Go to the user instructions for the selected key.

Contrasts

Object of Program

This program performs treatment contrasts on main effect means or on two-way means with one of the factors held constant. This allows you to make any desired linear contrast of a set of treatment means by entering an appropriate set of coefficients. The output includes the user-entered coefficients, the contrasts, and the sum of squares, F-ratio and P(F>comp F) associated with the contrasts.

NOTE

If the coefficients for the contrasts you enter are denoted by c_j and the sample sizes for the level of the factor are denoted by n_j, then one condition for choosing the c_j is that they must satisfy

$$\Sigma c_j n_j = 0 \quad (1)$$

where j is summed over all levels of the factor of interest. If $n_j = n$ for all levels, which is the case with several designs, then the condition merely states that the sum of the coefficients must be equal to zero. However, if one or more of the n_j are not the same, then the user must be careful when choosing the c_j to make sure they satisfy (1). Obviously, (1) implies that some of the c_j must be negative. Of course one or more of the c_j may be equal to zero.

User Instructions

1. When "Is the design displayed on the CRT the latest one?" is displayed:

   a. Press: NO if the treatment contrasts will not perform on the design displayed on the CRT.

   b. Go to step 15.

   or

   a. Press: YES if the design displayed on the CRT is the latest one.

2. For a one-way classification design, go to step 9.
3. When "Treatment contrasts on MAIN EFFECTS (enter 1) or TWO-WAY MEANS (enter 2)?" is displayed:

   a. Enter one if treatment contrasts on mean effects are desired.

   b. Press: CONT

   c. Go to step 4.

   or

   a. Enter two if treatment contrasts on two-way means are desired.

   b. Press: CONT

   c. Go to step 5.

4. When "Are treatment contrasts on factor (K) desired?" is displayed:

   a. Press: NO if treatment contrasts on factor (K) are not desired.

   b. Go to step 4.

   or

   a. Press: YES if treatment contrasts on factor (K) are desired.

   b. Go to step 9.

5. When "Are treatment contrasts (two-way mean) on factor (K) desired?" is displayed:

   a. Press: NO if treatment contrasts of two-way means on factor (K) are not desired.

   b. Go to step 5.

   or

   a. Press: YES if treatment contrasts of two-way means on factor (K) are desired.

6. For three-way, or four-way designs, go to step 8.
7. When "Enter the level number of Factor (L) to be held constant" is displayed:

   a. Enter the level number of Factor (L) which will be held constant in the two-way mean treatment contrasts.

   b. Press: CONT

8. When "Enter the name of factor (e.g. A,B) and level number, separated by a comma, to be held constant" is displayed:

   a. Enter the name of the factor and the level number to be held constant in treatment contrasts of the two-way means.

   b. Press: CONT

9. When "Enter coefficient 'T' " is displayed:

   a. Enter the contrast coefficient of treatment T.

   b. Press: CONT

   c. If more coefficients are to be entered, go to step 9.

10. When "Is the above information correct?" is displayed:

    a. Press: NO if changes are desired.

    b. Go to step 2.

    or

    a. Press: YES if no changes are necessary.

11. For the designs other than nested or split-plot, go to step 13.
12. When "Enter Error mean square and associated degrees of freedom" is displayed:
    a. Enter the error mean square and associated degrees of freedom to be used in this treatment contrast analysis.
    b. Press: CONT
13. The program will print out the corresponding contrast, sum of squares of contrasts, F ratio and corresponding probability.
14. When "Is another treatment contrast desired?" is displayed:
    a. Press: YES if another treatment contrast is desired.
    b. Go to step 2.

       or

    a. Press: NO if another treatment contrast is not desired.
15. When "SELECT ANY KEY" is displayed:
    a. Press any of the keys defined in the template.
    b. Go to the user instructions for the selected key.

Interaction Plots

Object of Program

This program will plot a two-way interaction, or a three-way interaction. The two-way interaction plot will be on one graph. You may decide which factor will be put on the X axis as well as the spacing of the levels, and then the other factor will be plotted. Each interaction line will be labeled indicating the level of the factor.

For instance, the three levels of a factor B will be labeled B1, B2, B3.

The three-way interaction plot will be plotted on several graphs. That is, a two-way interaction will be plotted for each level of the third factor. The program will give you a prompt when it is necessary to do the next page of the plot.

You may also have a legend drawn showing the length of the Least Significant Difference (LSD) and/or the length of Tukey's Honestly Significant Difference (HSD). To do these, it is necessary to enter the critical value, error mean square, and its corresponding degrees of freedom.

NOTE

When labeling the x-axis for an interaction plot, the factor levels must be given in increasing order. Factors whose levels are not in increasing order must be given arbitrary level codes if they are to be used on the x-axis of an interaction plot.

References

1. C.R. Hicks, "Fundamental Concepts in the Design of Experiments"; Second Edition. Holt, Rinehart, and Winston, 1972.
2. B.J. Winer, "Statistical Principles in Experimental Design"; Second Edition. McGraw-Hill, 1971.

User Instructions

1. The title will be printed.
2. If the last design key pressed was:

a. One Way AOV	go to step 4.
b. One Way AOC	go to step 4.
c. Factorial	go to step 5.
d. Nested	go to step 5.
e. Split Plot	go to step 5.
f. Unbalanced 2-way	go to step 5.
g. No design key has been pressed	go to step 3.

3. When "No design has been analyzed. Select another key." is displayed:

   a. This means that some design has to be specified before any plots can be made.

   b. Select a special function key and go to step 1 of the user instructions for the selected key.

4. When "No interactions exist for this design. Select another key." is displayed:

   a. The last design analyzed contained no interaction means and therefore cannot be plotted.

   b. Select a special function key and go to step 1 of the user instructions for the selected key.

5. A description of the current design is displayed.
6. When "Is this correct?" is displayed:

   a. Press: YES if this is the design you wish to analyze.

   b. Press: CONT

   c. Go to step 7.

   or

   a. Press: NO

   b. Go to step 36.

7. If only one two-way interaction is possible, go to step 14.
8. When "Is a two-way interaction to be plotted?" is displayed:

   a. Press: NO if no two-way interactions are to be plotted.

   b. Go to step 10.

   or

   c. Press: YES if a two-way interaction is to be plotted.

9. When "Which interaction is to be plotted: (list of possible choices)?" is displayed:

   a. Enter the two letters corresponding to the two factors. The input must be one of AB, AC, BC, AD, BD, or CD, and the one selected must be possible in the data set.

   b. Press: CONT

   c. Go to step 14.

10. If no three-way interactions are present, go to step 34.
11. If only one three-way interaction is present, go to step 13.
12. When "What interaction is to be plotted: (list of possible 3-way interactions)?" is displayed:

    a. Enter the three letters corresponding to the three factors. The input must be one of ABC, ABD, ACD, or BCD.

b. Press: CONT

13. When "Which factor is to be increased in levels across pages: (list of choices)?" is displayed:

    a. Enter the letter of the factor to be increased across plots. For each level of this factor, the other two factors will be plotted on one plot, then another plot will be made at the next level of this factor, and this will continue until all levels of this factor have been plotted. This will cause a plot to be made for each level of the factor.
    b. Press: CONT

14. When "Which factor is to be plotted on the X axis: (list of choices)?" is displayed:

    a. Enter the letter of the factor to be plotted on the X axis.
    b. Press: CONT

15. When "Enter the 'X' levels of Factor 'Y' separated by commas:?" is displayed:
    (X = no. of levels of factor Y   Y = factor specified in 14)

    a. Enter the levels separated by commas.
    b. Press: CONT

16. If there is only one response variable, go to step 18.
17. When "What is the name of the response? (<11 characters)" is displayed:

    a. Enter the name to be used to label the Y axis.
    b. Press: CONT

18. When "Enter Y minimum value. Must be less than 't'?" is displayed:

    a. Enter the minimum value to be plotted on the Y axis. This value must be smaller than the displayed value.
    b. Press: CONT

19. When "Enter Y maximum value. Must be greater than 't'?" is displayed:

    a. Enter the maximum value to be plotted on the Y axis. This value must be greater than the displayed value.
    b. Press: CONT

20. When "Enter Y tic" is displayed:

    a. Enter a positive value to be used as the spacing between tic marks on the Y axis.
    b. Press: CONT

21. When "Number of decimal places for labeling the Y axis?" is displayed:

    a. Enter a positive integer from 0 to 9.

    b. Press: CONT

22. When "Should the length of the LSD and/or HSD be plotted?" is displayed:

    a. Press NO if these lines should not be drawn.

    b. Go to step 24.

    or

    a. Press: YES if these line(s) will be plotted as a reference value on one side of the plot.

23. When "Enter the Error Mean Square to be used to calculate the LSD or HSD. Default values are: 'EMS Value,df' " is displayed:

    a. Enter the Error Mean Square and degrees of freedom separated by a comma.

    b. Press: CONT

    c. Go to step 24 if LSD or step 25 if HSD.

    or

    a. Press: CONT to use the default value.

    b. Go to step 24 if LSD or step 25 if HSD.

24. When "Enter the t value used for the LSD, or 0 to not plot the LSD" is displayed:

    a. Enter the t value and the length of the LSD will be drawn.

    b. Press: CONT

    c. Go to step 26.

    or

    a. Enter 0 for the t value and the LSD will not be drawn.

    b. Press: CONT

    c. Go to step 26.

25. When "Enter the Q value to use for the HSD, or 0 not to plot the HSD" is displayed:

    a. Enter the Q value and the length of the HSD will be drawn.

    b. Press: CONT

    c. Go to step 26.

    or

    a. Enter 0 for the Q value and the HSD will not be drawn.

    b. Press: CONT

*26. When "Should the 'plotter name' at select code 'sc' be used for plotting?" is displayed:

a. Press: YES if the displayed unit is correct.

b. Go to step 31.

or

a. Press: NO if another plotting unit is to be specified.

27. When "Enter the name of the plotter" is displayed:

a. Enter the name of the plotting unit. For example, GRAPHICS or 9872A.

b. Press: CONT

*28. If GRAPHICS was entered at step 27, go to step 31.

29. When "Enter the device select code" is displayed:

a. Enter the select code of the plotting device.

b. Press: CONT

30. When "Enter the device HB-IB address" is displayed:

a. Enter the device HB-IB address.

b. Press: CONT

31. When "Press CONT when the plotter is ready" is displayed:

a. When the plotter is ready, press CONT and a plot will be made.

*32. If the graphics display (CRT) was not used as a plotter, go to step 34.

*33. When "Do you want a hard copy?" is displayed:

a. Press: YES if the graphics display is to be dumped onto the thermal printer.

b. Go to step 34.

or

a. Press: NO if no hard copy of the plot is wanted.

34. If a three-way interaction plot is being made and further pages of the plot have to be made, go to step 31.

35. When "Are any more plots to be made?" is displayed:

a. Press: YES if another interaction plot is to be made.

b. Go to step 2.

* When using a 9835A, please delete these instructions: #26, 28, 32, 33.

or

a. Press: NO if all the interaction plots have been made.

36. When "SELECT KEY" is displayed, press any of the Special Function keys and go to step 1 of the user instructions for that key.

Multiple Comparisons

Object of Program

This program allows you to select any one of five multiple comparison procedures to use on either main effect means or two-way table means. You must input the appropriate tabled values for the procedure selected. In addition, for the separation procedures for the two-way means, you will need to specify the appropriate standard deviation to be used.

A separation table will be printed which should help you determine which treatment or factor levels are significantly different from one another. For example, the following table shows output for a set of treatments:

Factor A

Level	Mean.	Sample Size	Separation
1	10.7	10	ab
2	9.8	9	a
3	11.7	10	b
4	15.8	8	c

We would interpret this table as showing that factor level 4 is significantly different from the other levels of A since no other level has a "c" listed beside it. Also we see that level 1 cannot be distinguished from level 2 and level 1 cannot be distinguished from level 3. And, level 2 can be shown to significantly differ from level 3 since they have no letters in common.

Of course, the conclusion one draws from the separation procedure may depend on which procedure is used and the level of significance you choose.

References

1. Boardman, T.J. and D.R. Moffitt (1971) "Graphical Monte Carlo Type I Error for Multiple Comparison Procedures". Biometrics 27:3, 738-744.
2. Carmen, S.G., and M.R. Swanson (1973) "Evaluation of Ten Pairwise Multiple Comparison Procedures by Monte Carlo Methods". Journal of the American Statistical Association 68:341, pp. 66-74.
3. Duncan, D.B. (1955). Multiple range and multiple F tests. Biometrics 11, 1-42.
4. Pearson, E.S. and Hartley, H.O. (1958). Biometrika Tables for Statisticians, Vol. I. Cambridge University Press, London.
5. Scheffe,H. (1953). A method for judging all contrasts in the analysis of variance. Biometrika 40,87-104.
6. Tukey, J.W. (1953). The problem of multiple comparisons. Unpublished notes, Princeton University.

User Instructions

1. Press: Multiple Comparisons key.
2. If a one-way design was specified, go to step 5.
3. When "Enter 1 or 2 to specify Type of Means" is displayed:

   a. Enter 1 if multiple comparisons are to be determined on main effect means.
   b. Go to step 4.

   or

   a. Enter 2 if multiple comparisons are to be determined on two way table means.
   b. Go to step 25.

4. When "Which factor/main effect should be used?"

   a. Enter A, B, C, or D.
   b. Press: CONT

NOTE

Although factor labels may be used for labeling output, you should input A, B, C, or D for this response.

---

5. When "Which procedure would you like to use?" is displayed:

   a. Enter an integer from 0 to 5.
   b. Press: CONT
   c. If 0 was entered, go to step 33.

6. Depending on the procedure selected, the program will now ask you for the appropriate tabled value.
7. When "What level of alpha are you going to use?" is displayed:

   a. Enter the alpha level - usually a value between .01 and .10 is used.
   b. Press: CONT

NOTE

This value is used for printout purposes only and not for any calculations.

---

8. When "Enter (depends on procedure) tabled value with (parameters)" is displayed:

   a. Enter the tabled value.

   b. Press: CONT

The following table shows required inputs for tabled values:

Procedure#	Name	Notation	Parameter	Reference
1	Student's t	$t\alpha/2(df)$	df = error degrees of freedom	(1,4)
2	Studentized range	$q\alpha/2(p,df)$	p = # of means df = error degrees of freedom	(6,4)
3	Duncan's	$q^*\alpha(p,df)$	p is as above but reduces by 1 to p=2	(3)
4	Studentized range	$q\alpha/2(p,df)$	p same as 3	(1,3,4)
5	Snedecor's F	F(p-1, df)	p = # of means	(4,5)

* See references (1) and (2) for more information on all procedures.

9. If HSD or LSD multiple comparison procedures were not selected, go to step 19.

10. When "Is a plot of HSD(LSD) desired?" is displayed:

   a. Press: YES if a multiple comparison plot is desired. This plot includes the one-half LSD or HSD values above and below each factor level mean. If the sample sizes are unequal, then the harmonic mean sample size will be used.

   b. Go to step 11.

   or

   a. Press: NO if a plot is not desired.

   b. Go to step 19.

*11. When "Plot on CRT?" is displayed:

   a. Press: YES if the plot is to be shown on the CRT.

   b. Go to step 13.

   or

   a. Press: NO otherwise.

12. When "Enter plotter select code, HP bus #" is displayed:

   a. Enter the plotter select code and HP-IB device address separated by a comma.

   b. Press: CONT

13. When "Enter the name for labeling Y axis (<11 characters)" is displayed:

   a. Enter the name you wish to appear on the Y axis of the plot.

* When using a 9835A, delete this instruction and ignore any branches to it.

14. The LSD or HSD graph will now be plotted.

*15. When you have finished studying the plot:

   a. Press: CONT

*16. When "Do you want a hard copy of the plot?" is displayed:

   a. Press: YES if a hard copy is desired.
   b. Go to step 17.

   or

   a. Press: NO otherwise.

*17. When "Are you using perforated paper?" is displayed:

   a. Press: YES if you are using perforated paper.
   b. Go to step 19.

   or

   a. Press: NO otherwise.

*18. The plot is printed on paper.

19. When "Another separation procedure on factor (K)?" is displayed:

   a. Press: YES if you would like to choose one of the other multiple comparison procedures.
   b. Go to step 5.

   or

   a. Press: NO otherwise.

20. For a one-way design, go to step 33.

21. When "Another factor to be used?" is displayed:

   a. Press: YES if you would like to choose another factor on which to do the separation procedures.
   b. Go to step 4.

   or

   a. Press: NO otherwise.

22. The program checks to see if two-way interaction means are possible depending on the design previously chosen. If they are not, go to step 33.

* When using a 9835A, delete these instructions and ignore any branches to them: #15, 16, 17, 18.

23. When "Multiple comparison procedures on two-way means" is displayed:

    a. Press: YES if you would like to specify a two-way set of means on which you wish to perform multiple comparison procedures.
    b. Go to step 24.

    or

    a. Press: NO otherwise.
    b. When "SELECT KEY" is displayed,

    choose an appropriate key to further your analysis.

24. For a two-way design, go to step 26.
25. When "Which factors are to be studied?" is displayed:

    a. Enter the two factors (A,B,C,or D) separated by a comma.
    b. Press: CONT

26. When "Which factor will be held constant?" is displayed:

    a. Enter the factor (A,B,C or D) which will be held constant on the two-way means analysis.
    b. Press: CONT

27. When "Enter Error Mean Square and associated Degrees of Freedom" is displayed:

    a. Enter the appropriate error mean square and associated degrees of freedom to be used on the two-way means analysis.
    b. Press: CONT

28. When "Which procedure would you like to use?" is displayed:

    a. Enter an integer (from 0 to 5) depending on your choice.
    b. Press: CONT
    c. If 0 was entered, go to step 33.

29. Follow steps 6 through 18, then proceed to step 30.
30. When "Another separation procedure on Interaction Means of Factor (K) (L)?" is displayed:

    a. Press: YES if you would like to subject the (K) (L) interaction means to another multiple comparison procedure.
    b. Go to step 28.

31. If only two factors are in the design, go to step 33.

32. When "Choose another two-way interaction?" is displayed:

    a. Press: YES if you would like to specify another set of two-factor interaction means.

    b. Go to step 25.

    or

    a. Press: NO otherwise.

33. When "Select Key" is displayed:

    a. Choose any specific function key and go to the user instructions for that key.

Appendix

HP-IB

"HP-IB" stands for Hewlett-Packard Interface Bus and is our version of IEEE Standard 488,1975 which is a universal standard interface. In this program package it is used to refer to the interface between the computer and the 9872A Plotter. The select code may be found on the interface card. The bus address is found on the back of the plotter itself in binary form. For further information consult the Programming or Operating and Service Manual for the plotter.

Mass Storage Changes

To change the mass storage to a different tape drive, or to a disk, follow these instructions.

1. Type: MASS STORAGE IS ":unit specifier"
2. Press: EXECUTE

Some examples are ":T15" for the right hand side tape drive and ":F8" for the flexible disk master.

Notes

Notes

Notes

Percentage Points of the Duncan New Multiple Range Test

n_2 \ p	2	3	4	5	6	7	8	9	10	12	14	16	18	20	50	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.46	3.46	3.46	3.47	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.46	3.46	3.46	3.47	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.45	3.46	3.46	3.47	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.45	3.46	3.46	3.47	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.44	3.45	3.46	3.47	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.44	3.45	3.46	3.47	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.44	3.45	3.46	3.47	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.43	3.45	3.46	3.47	3.47	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.43	3.44	3.46	3.47	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.43	3.44	3.46	3.46	3.47	3.47	3.47
22	2.93	3.08	3.17	3.24	3.29	3.32	3.35	3.37	3.39	3.42	3.44	3.45	3.46	3.47	3.47	3.47
24	2.92	3.07	3.15	3.22	3.28	3.31	3.34	3.37	3.38	3.41	3.44	3.45	3.46	3.47	3.47	3.47
26	2.91	3.06	3.14	3.21	3.27	3.30	3.34	3.36	3.38	3.41	3.43	3.45	3.46	3.47	3.47	3.47
28	2.90	3.04	3.13	3.20	3.26	3.30	3.33	3.35	3.37	3.40	3.43	3.45	3.46	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.40	3.43	3.44	3.46	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.39	3.42	3.44	3.46	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.37	3.40	3.43	3.45	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.36	3.40	3.42	3.45	3.47	3.53	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.34	3.38	3.41	3.44	3.47	3.61	3.67

*Using special protection levels based on degrees of freedom.

Percentage Points of the Studentized Range, $q=(x_n-x_1)/s_\nu$.

Upper 10 % points

ν \ n	2	3	4	5	6	7	8	9	10
1	8·93	13·44	16·36	18·49	20·15	21·51	22·64	23·62	24·48
2	4·13	5·73	6·77	7·54	8·14	8·63	9·05	9·41	9·72
3	3·33	4·47	5·20	5·74	6·16	6·51	6·81	7·06	7·29
4	3·01	3·98	4·59	5·03	5·39	5·68	5·93	6·14	6·33
5	2·85	3·72	4·26	4·66	4·98	5·24	5·46	5·65	5·82
6	2·75	3·56	4·07	4·44	4·73	4·97	5·17	5·34	5·50
7	2·68	3·45	3·93	4·28	4·55	4·78	4·97	5·14	5·28
8	2·63	3·37	3·83	4·17	4·43	4·65	4·83	4·99	5·13
9	2·59	3·32	3·76	4·08	4·34	4·54	4·72	4·87	5·01
10	2·56	3·27	3·70	4·02	4·26	4·47	4·64	4·78	4·91
11	2·54	3·23	3·66	3·96	4·20	4·40	4·57	4·71	4·84
12	2·52	3·20	3·62	3·92	4·16	4·35	4·51	4·65	4·78
13	2·50	3·18	3·59	3·88	4·12	4·30	4·46	4·60	4·72
14	2·49	3·16	3·56	3·85	4·08	4·27	4·42	4·56	4·68
15	2·48	3·14	3·54	3·83	4·05	4·23	4·39	4·52	4·64
16	2·47	3·12	3·52	3·80	4·03	4·21	4·36	4·49	4·61
17	2·46	3·11	3·50	3·78	4·00	4·18	4·33	4·46	4·58
18	2·45	3·10	3·49	3·77	3·98	4·16	4·31	4·44	4·55
19	2·45	3·09	3·47	3·75	3·97	4·14	4·29	4·42	4·53
20	2·44	3·08	3·46	3·74	3·95	4·12	4·27	4·40	4·51
24	2·42	3·05	3·42	3·69	3·90	4·07	4·21	4·34	4·44
30	2·40	3·02	3·39	3·65	3·85	4·02	4·16	4·28	4·38
40	2·38	2·99	3·35	3·60	3·80	3·96	4·10	4·21	4·32
60	2·36	2·96	3·31	3·56	3·75	3·91	4·04	4·16	4·25
120	2·34	2·93	3·28	3·52	3·71	3·86	3·99	4·10	4·19
∞	2·33	2·90	3·24	3·48	3·66	3·81	3·93	4·04	4·13

ν \ n	11	12	13	14	15	16	17	18	19	20
1	25·24	25·92	26·54	27·10	27·62	28·10	28·54	28·96	29·35	29·71
2	10·01	10·26	10·49	10·70	10·89	11·07	11·24	11·39	11·54	11·68
3	7·49	7·67	7·83	7·98	8·12	8·25	8·37	8·48	8·58	8·68
4	6·49	6·65	6·78	6·91	7·02	7·13	7·23	7·33	7·41	7·50
5	5·97	6·10	6·22	6·34	6·44	6·54	6·63	6·71	6·79	6·86
6	5·64	5·76	5·87	5·98	6·07	6·16	6·25	6·32	6·40	6·47
7	5·41	5·53	5·64	5·74	5·83	5·91	5·99	6·06	6·13	6·19
8	5·25	5·36	5·46	5·56	5·64	5·72	5·80	5·87	5·93	6·00
9	5·13	5·23	5·33	5·42	5·51	5·58	5·66	5·72	5·79	5·85
10	5·03	5·13	5·23	5·32	5·40	5·47	5·54	5·61	5·67	5·73
11	4·95	5·05	5·15	5·23	5·31	5·38	5·45	5·51	5·57	5·63
12	4·89	4·99	5·08	5·16	5·24	5·31	5·37	5·44	5·49	5·55
13	4·83	4·93	5·02	5·10	5·18	5·25	5·31	5·37	5·43	5·48
14	4·79	4·88	4·97	5·05	5·12	5·19	5·26	5·32	5·37	5·43
15	4·75	4·84	4·93	5·01	5·08	5·15	5·21	5·27	5·32	5·38
16	4·71	4·81	4·89	4·97	5·04	5·11	5·17	5·23	5·28	5·33
17	4·68	4·77	4·86	4·93	5·01	5·07	5·13	5·19	5·24	5·30
18	4·65	4·75	4·83	4·90	4·98	5·04	5·10	5·16	5·21	5·26
19	4·63	4·72	4·80	4·88	4·95	5·01	5·07	5·13	5·18	5·23
20	4·61	4·70	4·78	4·85	4·92	4·99	5·05	5·10	5·16	5·20
24	4·54	4·63	4·71	4·78	4·85	4·91	4·97	5·02	5·07	5·12
30	4·47	4·56	4·64	4·71	4·77	4·83	4·89	4·94	4·99	5·03
40	4·41	4·49	4·56	4·63	4·69	4·75	4·81	4·86	4·90	4·95
60	4·34	4·42	4·49	4·56	4·62	4·67	4·73	4·78	4·82	4·86
120	4·28	4·35	4·42	4·48	4·54	4·60	4·65	4·69	4·74	4·78
∞	4·21	4·28	4·35	4·41	4·47	4·52	4·57	4·61	4·65	4·69

n: size of sample from which range obtained. ν: degrees of freedom of independent s_ν.

Percentage Points of the Studentized Range, $q=(x_n-x_1)/s_\nu$. (continued)

Upper 5% points

ν \ n	2	3	4	5	6	7	8	9	10
1	17·97	26·98	32·82	37·08	40·41	43·12	45·40	47·36	49·07
2	6·08	8·33	9·80	10·88	11·74	12·44	13·03	13·54	13·99
3	4·50	5·91	6·82	7·50	8·04	8·48	8·85	9·18	9·46
4	3·93	5·04	5·76	6·29	6·71	7·05	7·35	7·60	7·83
5	3·64	4·60	5·22	5·67	6·03	6·33	6·58	6·80	6·99
6	3·46	4·34	4·90	5·30	5·63	5·90	6·12	6·32	6·49
7	3·34	4·16	4·68	5·06	5·36	5·61	5·82	6·00	6·16
8	3·26	4·04	4·53	4·89	5·17	5·40	5·60	5·77	5·92
9	3·20	3·95	4·41	4·76	5·02	5·24	5·43	5·59	5·74
10	3·15	3·88	4·33	4·65	4·91	5·12	5·30	5·46	5·60
11	3·11	3·82	4·26	4·57	4·82	5·03	5·20	5·35	5·49
12	3·08	3·77	4·20	4·51	4·75	4·95	5·12	5·27	5·39
13	3·06	3·73	4·15	4·45	4·69	4·88	5·05	5·19	5·32
14	3·03	3·70	4·11	4·41	4·64	4·83	4·99	5·13	5·25
15	3·01	3·67	4·08	4·37	4·59	4·78	4·94	5·08	5·20
16	3·00	3·65	4·05	4·33	4·56	4·74	4·90	5·03	5·15
17	2·98	3·63	4·02	4·30	4·52	4·70	4·86	4·99	5·11
18	2·97	3·61	4·00	4·28	4·49	4·67	4·82	4·96	5·07
19	2·96	3·59	3·98	4·25	4·47	4·65	4·79	4·92	5·04
20	2·95	3·58	3·96	4·23	4·45	4·62	4·77	4·90	5·01
24	2·92	3·53	3·90	4·17	4·37	4·54	4·68	4·81	4·92
30	2·89	3·49	3·85	4·10	4·30	4·46	4·60	4·72	4·82
40	2·86	3·44	3·79	4·04	4·23	4·39	4·52	4·63	4·73
60	2·83	3·40	3·74	3·98	4·16	4·31	4·44	4·55	4·65
120	2·80	3·36	3·68	3·92	4·10	4·24	4·36	4·47	4·56
∞	2·77	3·31	3·63	3·86	4·03	4·17	4·29	4·39	4·47

ν \ n	11	12	13	14	15	16	17	18	19	20
1	50·59	51·96	53·20	54·33	55·36	56·32	57·22	58·04	58·83	59·56
2	14·39	14·75	15·08	15·38	15·65	15·91	16·14	16·37	16·57	16·77
3	9·72	9·95	10·15	10·35	10·52	10·69	10·84	10·98	11·11	11·24
4	8·03	8·21	8·37	8·52	8·66	8·79	8·91	9·03	9·13	9·23
5	7·17	7·32	7·47	7·60	7·72	7·83	7·93	8·03	8·12	8·21
6	6·65	6·79	6·92	7·03	7·14	7·24	7·34	7·43	7·51	7·59
7	6·30	6·43	6·55	6·66	6·76	6·85	6·94	7·02	7·10	7·17
8	6·05	6·18	6·29	6·39	6·48	6·57	6·65	6·73	6·80	6·87
9	5·87	5·98	6·09	6·19	6·28	6·36	6·44	6·51	6·58	6·64
10	5·72	5·83	5·93	6·03	6·11	6·19	6·27	6·34	6·40	6·47
11	5·61	5·71	5·81	5·90	5·98	6·06	6·13	6·20	6·27	6·33
12	5·51	5·61	5·71	5·80	5·88	5·95	6·02	6·09	6·15	6·21
13	5·43	5·53	5·63	5·71	5·79	5·86	5·93	5·99	6·05	6·11
14	5·36	5·46	5·55	5·64	5·71	5·79	5·85	5·91	5·97	6·03
15	5·31	5·40	5·49	5·57	5·65	5·72	5·78	5·85	5·90	5·96
16	5·26	5·35	5·44	5·52	5·59	5·66	5·73	5·79	5·84	5·90
17	5·21	5·31	5·39	5·47	5·54	5·61	5·67	5·73	5·79	5·84
18	5·17	5·27	5·35	5·43	5·50	5·57	5·63	5·69	5·74	5·79
19	5·14	5·23	5·31	5·39	5·46	5·53	5·59	5·65	5·70	5·75
20	5·11	5·20	5·28	5·36	5·43	5·49	5·55	5·61	5·66	5·71
24	5·01	5·10	5·18	5·25	5·32	5·38	5·44	5·49	5·55	5·59
30	4·92	5·00	5·08	5·15	5·21	5·27	5·33	5·38	5·43	5·47
40	4·82	4·90	4·98	5·04	5·11	5·16	5·22	5·27	5·31	5·36
60	4·73	4·81	4·88	4·94	5·00	5·06	5·11	5·15	5·20	5·24
120	4·64	4·71	4·78	4·84	4·90	4·95	5·00	5·04	5·09	5·13
∞	4·55	4·62	4·68	4·74	4·80	4·85	4·89	4·93	4·97	5·01

n: size of sample from which range obtained. ν: degrees of freedom of independent s_ν.

Percentage Points of the Studentized Range, $q=(x_n-x_1)/s_\nu$. (continued)

Upper 1% points

ν \ n	2	3	4	5	6	7	8	9	10
1	90·03	135·0	164·3	185·6	202·2	215·8	227·2	237·0	245·6
2	14·04	19·02	22·29	24·72	26·63	28·20	29·53	30·68	31·69
3	8·26	10·62	12·17	13·33	14·24	15·00	15·64	16·20	16·69
4	6·51	8·12	9·17	9·96	10·58	11·10	11·55	11·93	12·27
5	5·70	6·98	7·80	8·42	8·91	9·32	9·67	9·97	10·24
6	5·24	6·33	7·03	7·56	7·97	8·32	8·61	8·87	9·10
7	4·95	5·92	6·54	7·01	7·37	7·68	7·94	8·17	8·37
8	4·75	5·64	6·20	6·62	6·96	7·24	7·47	7·68	7·86
9	4·60	5·43	5·96	6·35	6·66	6·91	7·13	7·33	7·49
10	4·48	5·27	5·77	6·14	6·43	6·67	6·87	7·05	7·21
11	4·39	5·15	5·62	5·97	6·25	6·48	6·67	6·84	6·99
12	4·32	5·05	5·50	5·84	6·10	6·32	6·51	6·67	6·81
13	4·26	4·96	5·40	5·73	5·98	6·19	6·37	6·53	6·67
14	4·21	4·89	5·32	5·63	5·88	6·08	6·26	6·41	6·54
15	4·17	4·84	5·25	5·56	5·80	5·99	6·16	6·31	6·44
16	4·13	4·79	5·19	5·49	5·72	5·92	6·08	6·22	6·35
17	4·10	4·74	5·14	5·43	5·66	5·85	6·01	6·15	6·27
18	4·07	4·70	5·09	5·38	5·60	5·79	5·94	6·08	6·20
19	4·05	4·67	5·05	5·33	5·55	5·73	5·89	6·02	6·14
20	4·02	4·64	5·02	5·29	5·51	5·69	5·84	5·97	6·09
24	3·96	4·55	4·91	5·17	5·37	5·54	5·69	5·81	5·92
30	3·89	4·45	4·80	5·05	5·24	5·40	5·54	5·65	5·76
40	3·82	4·37	4·70	4·93	5·11	5·26	5·39	5·50	5·60
60	3·76	4·28	4·59	4·82	4·99	5·13	5·25	5·36	5·45
120	3·70	4·20	4·50	4·71	4·87	5·01	5·12	5·21	5·30
∞	3·64	4·12	4·40	4·60	4·76	4·88	4·99	5·08	5·16

ν \ n	11	12	13	14	15	16	17	18	19	20
1	253·2	260·0	266·2	271·8	277·0	281·8	286·3	290·4	294·3	298·0
2	32·59	33·40	34·13	34·81	35·43	36·00	36·53	·37·03	37·50	37·95
3	17·13	17·53	17·89	18·22	18·52	18·81	19·07	19·32	19·55	19·77
4	12·57	12·84	13·09	13·32	13·53	13·73	13·91	14·08	14·24	14·40
5	10·48	10·70	10·89	11·08	11·24	11·40	11·55	11·68	11·81	11·93
6	9·30	9·48	9·65	9·81	9·95	10·08	10·21	10·32	10·43	10·54
7	8·55	8·71	8·86	9·00	9·12	9·24	9·35	9·46	9·55	9·65
8	8·03	8·18	8·31	8·44	8·55	8·66	8·76	8·85	8·94	9·03
9	7·65	7·78	7·91	8·03	8·13	8·23	8·33	8·41	8·49	8·57
10	7·36	7·49	7·60	7·71	7·81	7·91	7·99	8·08	8·15	8·23
11	7·13	7·25	7·36	7·46	7·56	7·65	7·73	7·81	7·88	7·95
12	6·94	7·06	7·17	7·26	7·36	7·44	7·52	7·59	7·66	7·73
13	6·79	6·90	7·01	7·10	7·19	7·27	7·35	7·42	7·48	7·55
14	6·66	6·77	6·87	6·96	7·05	7·13	7·20	7·27	7·33	7·39
15	6·55	6·66	6·76	6·84	6·93	7·00	7·07	7·14	7·20	7·26
16	6·46	6·56	6·66	6·74	6·82	6·90	6·97	7·03	7·09	7·15
17	6·38	6·48	6·57	6·66	6·73	6·81	6·87	6·94	7·00	7·05
18	6·31	6·41	6·50	6·58	6·65	6·73	6·79	6·85	6·91	6·97
19	6·25	6·34	6·43	6·51	6·58	6·65	6·72	6·78	6·84	6·89
20	6·19	6·28	6·37	6·45	6·52	6·59	6·65	6·71	6·77	6·82
24	6·02	6·11	6·19	6·26	6·33	6·39	6·45	6·51	6·56	6·61
30	5·85	5·93	6·01	6·08	6·14	6·20	6·26	6·31	6·36	6·41
40	5·69	5·76	5·83	5·90	5·96	6·02	6·07	6·12	6·16	6·21
60	5·53	5·60	5·67	5·73	5·78	5·84	5·89	5·93	5·97	6·01
120	5·37	5·44	5·50	5·56	5·61	5·66	5·71	5·75	5·79	5·83
∞	5·23	5·29	5·35	5·40	5·45	5·49	5·54	5·57	5·61	5·65

Percentage Points of the F-distribution (Variance Ratio)

Upper 25% points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	5·83	7·50	8·20	8·58	8·82	8·98	9·10	9·19	9·26	9·32	9·41	9·49	9·58	9·63	9·67	9·71	9·76	9·80	9·85
2	2·57	3·00	3·15	3·23	3·28	3·31	3·34	3·35	3·37	3·38	3·39	3·41	3·43	3·43	3·44	3·45	3·46	3·47	3·48
3	2·02	2·28	2·36	2·39	2·41	2·42	2·43	2·44	2·44	2·44	2·45	2·46	2·46	2·46	2·47	2·47	2·47	2·47	2·47
4	1·81	2·00	2·05	2·06	2·07	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08	2·08
5	1·69	1·85	1·88	1·89	1·89	1·89	1·89	1·89	1·89	1·89	1·89	1·89	1·88	1·88	1·88	1·88	1·87	1·87	1·87
6	1·62	1·76	1·78	1·79	1·79	1·78	1·78	1·78	1·77	1·77	1·77	1·76	1·76	1·75	1·75	1·75	1·74	1·74	1·74
7	1·57	1·70	1·72	1·72	1·71	1·71	1·70	1·70	1·69	1·69	1·68	1·68	1·67	1·67	1·66	1·66	1·65	1·65	1·65
8	1·54	1·66	1·67	1·66	1·66	1·65	1·64	1·64	1·63	1·63	1·62	1·62	1·61	1·60	1·60	1·59	1·59	1·58	1·58
9	1·51	1·62	1·63	1·63	1·62	1·61	1·60	1·60	1·59	1·59	1·58	1·57	1·56	1·56	1·55	1·54	1·54	1·53	1·53
10	1·49	1·60	1·60	1·59	1·59	1·58	1·57	1·56	1·56	1·55	1·54	1·53	1·52	1·52	1·51	1·51	1·50	1·49	1·48
11	1·47	1·58	1·58	1·57	1·56	1·55	1·54	1·53	1·53	1·52	1·51	1·50	1·49	1·49	1·48	1·47	1·47	1·46	1·45
12	1·46	1·56	1·56	1·55	1·54	1·53	1·52	1·51	1·51	1·50	1·49	1·48	1·47	1·46	1·45	1·45	1·44	1·43	1·42
13	1·45	1·55	1·55	1·53	1·52	1·51	1·50	1·49	1·49	1·48	1·47	1·46	1·45	1·44	1·43	1·42	1·42	1·41	1·40
14	1·44	1·53	1·53	1·52	1·51	1·50	1·49	1·48	1·47	1·46	1·45	1·44	1·43	1·42	1·41	1·41	1·40	1·39	1·38
15	1·43	1·52	1·52	1·51	1·49	1·48	1·47	1·46	1·46	1·45	1·44	1·43	1·41	1·41	1·40	1·39	1·38	1·37	1·36
16	1·42	1·51	1·51	1·50	1·48	1·47	1·46	1·45	1·44	1·44	1·43	1·41	1·40	1·39	1·38	1·37	1·36	1·35	1·34
17	1·42	1·51	1·50	1·49	1·47	1·46	1·45	1·44	1·43	1·43	1·41	1·40	1·39	1·38	1·37	1·36	1·35	1·34	1·33
18	1·41	1·50	1·49	1·48	1·46	1·45	1·44	1·43	1·42	1·42	1·40	1·39	1·38	1·37	1·36	1·35	1·34	1·33	1·32
19	1·41	1·49	1·49	1·47	1·46	1·44	1·43	1·42	1·41	1·41	1·40	1·38	1·37	1·36	1·35	1·34	1·33	1·32	1·30
20	1·40	1·49	1·48	1·47	1·45	1·44	1·43	1·42	1·41	1·40	1·39	1·37	1·36	1·35	1·34	1·33	1·32	1·31	1·29
21	1·40	1·48	1·48	1·46	1·44	1·43	1·42	1·41	1·40	1·39	1·38	1·37	1·35	1·34	1·33	1·32	1·31	1·30	1·28
22	1·40	1·48	1·47	1·45	1·44	1·42	1·41	1·40	1·39	1·39	1·37	1·36	1·34	1·33	1·32	1·31	1·30	1·29	1·28
23	1·39	1·47	1·47	1·45	1·43	1·42	1·41	1·40	1·39	1·38	1·37	1·35	1·34	1·33	1·32	1·31	1·30	1·28	1·27
24	1·39	1·47	1·46	1·44	1·43	1·41	1·40	1·39	1·38	1·38	1·36	1·35	1·33	1·32	1·31	1·30	1·29	1·28	1·26
25	1·39	1·47	1·46	1·44	1·42	1·41	1·40	1·39	1·38	1·37	1·36	1·34	1·33	1·32	1·31	1·29	1·28	1·27	1·25
26	1·38	1·46	1·45	1·44	1·42	1·41	1·39	1·38	1·37	1·37	1·35	1·34	1·32	1·31	1·30	1·29	1·28	1·26	1·25
27	1·38	1·46	1·45	1·43	1·42	1·40	1·39	1·38	1·37	1·36	1·35	1·33	1·32	1·31	1·30	1·28	1·27	1·26	1·24
28	1·38	1·46	1·45	1·43	1·41	1·40	1·39	1·38	1·37	1·36	1·34	1·33	1·31	1·30	1·29	1·28	1·27	1·25	1·24
29	1·38	1·45	1·45	1·43	1·41	1·40	1·38	1·37	1·36	1·35	1·34	1·32	1·31	1·30	1·29	1·27	1·26	1·25	1·23
30	1·38	1·45	1·44	1·42	1·41	1·39	1·38	1·37	1·36	1·35	1·34	1·32	1·30	1·29	1·28	1·27	1·26	1·24	1·23
40	1·36	1·44	1·42	1·40	1·39	1·37	1·36	1·35	1·34	1·33	1·31	1·30	1·28	1·26	1·25	1·24	1·22	1·21	1·19
60	1·35	1·42	1·41	1·38	1·37	1·35	1·33	1·32	1·31	1·30	1·29	1·27	1·25	1·24	1·22	1·21	1·19	1·17	1·15
120	1·34	1·40	1·39	1·37	1·35	1·33	1·31	1·30	1·29	1·28	1·26	1·24	1·22	1·21	1·19	1·18	1·16	1·13	1·10
∞	1·32	1·39	1·37	1·35	1·33	1·31	1·29	1·28	1·27	1·25	1·24	1·22	1·19	1·18	1·16	1·14	1·12	1·08	1·00

$F=\dfrac{s_1^2}{s_2^2}=\dfrac{S_1}{\nu_1}\Big/\dfrac{S_2}{\nu_2}$, where $s_1^2=S_1/\nu_1$ and $s_2^2=S_2/\nu_2$ are independent mean squares estimating a common variance σ^2 and based on ν_1 and ν_2 degrees of freedom, respectively.

Percentage Points of the F-distribution (Variance Ratio) (continued)

Upper 10% points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	59·86	60·19	60·71	61·22	61·74	62·00	62·26	62·53	62·79	63·06	63·33
2	8·53	9·00	9·16	9·24	9·29	9·33	9·35	9·37	9·38	9·39	9·41	9·42	9·44	9·45	9·46	9·47	9·47	9·48	9·49
3	5·54	5·46	5·39	5·34	5·31	5·28	5·27	5·25	5·24	5·23	5·22	5·20	5·18	5·18	5·17	5·16	5·15	5·14	5·13
4	4·54	4·32	4·19	4·11	4·05	4·01	3·98	3·95	3·94	3·92	3·90	3·87	3·84	3·83	3·82	3·80	3·79	3·78	3·76
5	4·06	3·78	3·62	3·52	3·45	3·40	3·37	3·34	3·32	3·30	3·27	3·24	3·21	3·19	3·17	3·16	3·14	3·12	3·10
6	3·78	3·46	3·29	3·18	3·11	3·05	3·01	2·98	2·96	2·94	2·90	2·87	2·84	2·82	2·80	2·78	2·76	2·74	2·72
7	3·59	3·26	3·07	2·96	2·88	2·83	2·78	2·75	2·72	2·70	2·67	2·63	2·59	2·58	2·56	2·54	2·51	2·49	2·47
8	3·46	3·11	2·92	2·81	2·73	2·67	2·62	2·59	2·56	2·54	2·50	2·46	2·42	2·40	2·38	2·36	2·34	2·32	2·29
9	3·36	3·01	2·81	2·69	2·61	2·55	2·51	2·47	2·44	2·42	2·38	2·34	2·30	2·28	2·25	2·23	2·21	2·18	2·16
10	3·29	2·92	2·73	2·61	2·52	2·46	2·41	2·38	2·35	2·32	2·28	2·24	2·20	2·18	2·16	2·13	2·11	2·08	2·06
11	3·23	2·86	2·66	2·54	2·45	2·39	2·34	2·30	2·27	2·25	2·21	2·17	2·12	2·10	2·08	2·05	2·03	2·00	1·97
12	3·18	2·81	2·61	2·48	2·39	2·33	2·28	2·24	2·21	2·19	2·15	2·10	2·06	2·04	2·01	1·99	1·96	1·93	1·90
13	3·14	2·76	2·56	2·43	2·35	2·28	2·23	2·20	2·16	2·14	2·10	2·05	2·01	1·98	1·96	1·93	1·90	1·88	1·85
14	3·10	2·73	2·52	2·39	2·31	2·24	2·19	2·15	2·12	2·10	2·05	2·01	1·96	1·94	1·91	1·89	1·86	1·83	1·80
15	3·07	2·70	2·49	2·36	2·27	2·21	2·16	2·12	2·09	2·06	2·02	1·97	1·92	1·90	1·87	1·85	1·82	1·79	1·76
16	3·05	2·67	2·46	2·33	2·24	2·18	2·13	2·09	2·06	2·03	1·99	1·94	1·89	1·87	1·84	1·81	1·78	1·75	1·72
17	3·03	2·64	2·44	2·31	2·22	2·15	2·10	2·06	2·03	2·00	1·96	1·91	1·86	1·84	1·81	1·78	1·75	1·72	1·69
18	3·01	2·62	2·42	2·29	2·20	2·13	2·08	2·04	2·00	1·98	1·93	1·89	1·84	1·81	1·78	1·75	1·72	1·69	1·66
19	2·99	2·61	2·40	2·27	2·18	2·11	2·06	2·02	1·98	1·96	1·91	1·86	1·81	1·79	1·76	1·73	1·70	1·67	1·63
20	2·97	2·59	2·38	2·25	2·16	2·09	2·04	2·00	1·96	1·94	1·89	1·84	1·79	1·77	1·74	1·71	1·68	1·64	1·61
21	2·96	2·57	2·36	2·23	2·14	2·08	2·02	1·98	1·95	1·92	1·87	1·83	1·78	1·75	1·72	1·69	1·66	1·62	1·59
22	2·95	2·56	2·35	2·22	2·13	2·06	2·01	1·97	1·93	1·90	1·86	1·81	1·76	1·73	1·70	1·67	1·64	1·60	1·57
23	2·94	2·55	2·34	2·21	2·11	2·05	1·99	1·95	1·92	1·89	1·84	1·80	1·74	1·72	1·69	1·66	1·62	1·59	1·55
24	2·93	2·54	2·33	2·19	2·10	2·04	1·98	1·94	1·91	1·88	1·83	1·78	1·73	1·70	1·67	1·64	1·61	1·57	1·53
25	2·92	2·53	2·32	2·18	2·09	2·02	1·97	1·93	1·89	1·87	1·82	1·77	1·72	1·69	1·66	1·63	1·59	1·56	1·52
26	2·91	2·52	2·31	2·17	2·08	2·01	1·96	1·92	1·88	1·86	1·81	1·76	1·71	1·68	1·65	1·61	1·58	1·54	1·50
27	2·90	2·51	2·30	2·17	2·07	2·00	1·95	1·91	1·87	1·85	1·80	1·75	1·70	1·67	1·64	1·60	1·57	1·53	1·49
28	2·89	2·50	2·29	2·16	2·06	2·00	1·94	1·90	1·87	1·84	1·79	1·74	1·69	1·66	1·63	1·59	1·56	1·52	1·48
29	2·89	2·50	2·28	2·15	2·06	1·99	1·93	1·89	1·86	1·83	1·78	1·73	1·68	1·65	1·62	1·58	1·55	1·51	1·47
30	2·88	2·49	2·28	2·14	2·05	1·98	1·93	1·88	1·85	1·82	1·77	1·72	1·67	1·64	1·61	1·57	1·54	1·50	1·46
40	2·84	2·44	2·23	2·09	2·00	1·93	1·87	1·83	1·79	1·76	1·71	1·66	1·61	1·57	1·54	1·51	1·47	1·42	1·38
60	2·79	2·39	2·18	2·04	1·95	1·87	1·82	1·77	1·74	1·71	1·66	1·60	1·54	1·51	1·48	1·44	1·40	1·35	1·29
120	2·75	2·35	2·13	1·99	1·90	1·82	1·77	1·72	1·68	1·65	1·60	1·55	1·48	1·45	1·41	1·37	1·32	1·26	1·19
∞	2·71	2·30	2·08	1·94	1·85	1·77	1·72	1·67	1·63	1·60	1·55	1·49	1·42	1·38	1·34	1·30	1·24	1·17	1·00

$F = \frac{s_1^2}{s_2^2} = \frac{S_1}{\nu_1} \Big/ \frac{S_2}{\nu_2}$, where $s_1^2 = S_1/\nu_1$ and $s_2^2 = S_2/\nu_2$ are independent mean squares estimating a common variance σ^2 and based on ν_1 and ν_2 degrees of freedom, respectively.

Percentage Points of the F-distribution (Variance Ratio) (continued)

Upper 5 % points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	240·5	241·9	243·9	245·9	248·0	249·1	250·1	251·1	252·2	253·3	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·38	19·40	19·41	19·43	19·45	19·45	19·46	19·47	19·48	19·49	19·50
3	10·13	9·55	9·28	9·12	9·01	8·94	8·89	8·85	8·81	8·79	8·74	8·70	8·66	8·64	8·62	8·59	8·57	8·55	8·53
4	7·71	6·94	6·59	6·39	6·26	6·16	6·09	6·04	6·00	5·96	5·91	5·86	5·80	5·77	5·75	5·72	5·69	5·66	5·63
5	6·61	5·79	5·41	5·19	5·05	4·95	4·88	4·82	4·77	4·74	4·68	4·62	4·56	4·53	4·50	4·46	4·43	4·40	4·36
6	5·99	5·14	4·76	4·53	4·39	4·28	4·21	4·15	4·10	4·06	4·00	3·94	3·87	3·84	3·81	3·77	3·74	3·70	3·67
7	5·59	4·74	4·35	4·12	3·97	3·87	3·79	3·73	3·68	3·64	3·57	3·51	3·44	3·41	3·38	3·34	3·30	3·27	3·23
8	5·32	4·46	4·07	3·84	3·69	3·58	3·50	3·44	3·39	3·35	3·28	3·22	3·15	3·12	3·08	3·04	3·01	2·97	2·93
9	5·12	4·26	3·86	3·63	3·48	3·37	3·29	3·23	3·18	3·14	3·07	3·01	2·94	2·90	2·86	2·83	2·79	2·75	2·71
10	4·96	4·10	3·71	3·48	3·33	3·22	3·14	3·07	3·02	2·98	2·91	2·85	2·77	2·74	2·70	2·66	2·62	2·58	2·54
11	4·84	3·98	3·59	3·36	3·20	3·09	3·01	2·95	2·90	2·85	2·79	2·72	2·65	2·61	2·57	2·53	2·49	2·45	2·40
12	4·75	3·89	3·49	3·26	3·11	3·00	2·91	2·85	2·80	2·75	2·69	2·62	2·54	2·51	2·47	2·43	2·38	2·34	2·30
13	4·67	3·81	3·41	3·18	3·03	2·92	2·83	2·77	2·71	2·67	2·60	2·53	2·46	2·42	2·38	2·34	2·30	2·25	2·21
14	4·60	3·74	3·34	3·11	2·96	2·85	2·76	2·70	2·65	2·60	2·53	2·46	2·39	2·35	2·31	2·27	2·22	2·18	2·13
15	4·54	3·68	3·29	3·06	2·90	2·79	2·71	2·64	2·59	2·54	2·48	2·40	2·33	2·29	2·25	2·20	2·16	2·11	2·07
16	4·49	3·63	3·24	3·01	2·85	2·74	2·66	2·59	2·54	2·49	2·42	2·35	2·28	2·24	2·19	2·15	2·11	2·06	2·01
17	4·45	3·59	3·20	2·96	2·81	2·70	2·61	2·55	2·49	2·45	2·38	2·31	2·23	2·19	2·15	2·10	2·06	2·01	1·96
18	4·41	3·55	3·16	2·93	2·77	2·66	2·58	2·51	2·46	2·41	2·34	2·27	2·19	2·15	2·11	2·06	2·02	1·97	1·92
19	4·38	3·52	3·13	2·90	2·74	2·63	2·54	2·48	2·42	2·38	2·31	2·23	2·16	2·11	2·07	2·03	1·98	1·93	1·88
20	4·35	3·49	3·10	2·87	2·71	2·60	2·51	2·45	2·39	2·35	2·28	2·20	2·12	2·08	2·04	1·99	1·95	1·90	1·84
21	4·32	3·47	3·07	2·84	2·68	2·57	2·49	2·42	2·37	2·32	2·25	2·18	2·10	2·05	2·01	1·96	1·92	1·87	1·81
22	4·30	3·44	3·05	2·82	2·66	2·55	2·46	2·40	2·34	2·30	2·23	2·15	2·07	2·03	1·98	1·94	1·89	1·84	1·78
23	4·28	3·42	3·03	2·80	2·64	2·53	2·44	2·37	2·32	2·27	2·20	2·13	2·05	2·01	1·96	1·91	1·86	1·81	1·76
24	4·26	3·40	3·01	2·78	2·62	2·51	2·42	2·36	2·30	2·25	2·18	2·11	2·03	1·98	1·94	1·89	1·84	1·79	1·73
25	4·24	3·39	2·99	2·76	2·60	2·49	2·40	2·34	2·28	2·24	2·16	2·09	2·01	1·96	1·92	1·87	1·82	1·77	1·71
26	4·23	3·37	2·98	2·74	2·59	2·47	2·39	2·32	2·27	2·22	2·15	2·07	1·99	1·95	1·90	1·85	1·80	1·75	1·69
27	4·21	3·35	2·96	2·73	2·57	2·46	2·37	2·31	2·25	2·20	2·13	2·06	1·97	1·93	1·88	1·84	1·79	1·73	1·67
28	4·20	3·34	2·95	2·71	2·56	2·45	2·36	2·29	2·24	2·19	2·12	2·04	1·96	1·91	1·87	1·82	1·77	1·71	1·65
29	4·18	3·33	2·93	2·70	2·55	2·43	2·35	2·28	2·22	2·18	2·10	2·03	1·94	1·90	1·85	1·81	1·75	1·70	1·64
30	4·17	3·32	2·92	2·69	2·53	2·42	2·33	2·27	2·21	2·16	2·09	2·01	1·93	1·89	1·84	1·79	1·74	1·68	1·62
40	4·08	3·23	2·84	2·61	2·45	2·34	2·25	2·18	2·12	2·08	2·00	1·92	1·84	1·79	1·74	1·69	1·64	1·58	1·51
60	4·00	3·15	2·76	2·53	2·37	2·25	2·17	2·10	2·04	1·99	1·92	1·84	1·75	1·70	1·65	1·59	1·53	1·47	1·39
120	3·92	3·07	2·68	2·45	2·29	2·17	2·09	2·02	1·96	1·91	1·83	1·75	1·66	1·61	1·55	1·50	1·43	1·35	1·25
∞	3·84	3·00	2·60	2·37	2·21	2·10	2·01	1·94	1·88	1·83	1·75	1·67	1·57	1·52	1·46	1·39	1·32	1·22	1·00

$F = \frac{s_1^2}{s_2^2} = \frac{S_1}{\nu_1} \Big/ \frac{S_2}{\nu_2}$, where $s_1^2 = S_1/\nu_1$ and $s_2^2 = S_2/\nu_2$ are independent mean squares estimating a common variance σ^2 and based on ν_1 and ν_2 degrees of freedom, respectively.

Percentage Points of the F-distribution (Variance Ratio) (continued)

Upper 0·5 % points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	24940	25044	25148	25253	25359	25465
2	198·5	199·0	199·2	199·2	199·3	199·3	199·4	199·4	199·4	199·4	199·4	199·4	199·4	199·5	199·5	199·5	199·5	199·5	199·5
3	55·55	49·80	47·47	46·19	45·39	44·84	44·43	44·13	43·88	43·69	43·39	43·08	42·78	42·62	42·47	42·31	42·15	41·99	41·83
4	31·33	26·28	24·26	23·15	22·46	21·97	21·62	21·35	21·14	20·97	20·70	20·44	20·17	20·03	19·89	19·75	19·61	19·47	19·32
5	22·78	18·31	16·53	15·56	14·94	14·51	14·20	13·96	13·77	13·62	13·38	13·15	12·90	12·78	12·66	12·53	12·40	12·27	12·14
6	18·63	14·54	12·92	12·03	11·46	11·07	10·79	10·57	10·39	10·25	10·03	9·81	9·59	9·47	9·36	9·24	9·12	9·00	8·88
7	16·24	12·40	10·88	10·05	9·52	9·16	8·89	8·68	8·51	8·38	8·18	7·97	7·75	7·65	7·53	7·42	7·31	7·19	7·08
8	14·69	11·04	9·60	8·81	8·30	7·95	7·69	7·50	7·34	7·21	7·01	6·81	6·61	6·50	6·40	6·29	6·18	6·06	5·95
9	13·61	10·11	8·72	7·96	7·47	7·13	6·88	6·69	6·54	6·42	6·23	6·03	5·83	5·73	5·62	5·52	5·41	5·30	5·19
10	12·83	9·43	8·08	7·34	6·87	6·54	6·30	6·12	5·97	5·85	5·66	5·47	5·27	5·17	5·07	4·97	4·86	4·75	4·64
11	12·23	8·91	7·60	6·88	6·42	6·10	5·86	5·68	5·54	5·42	5·24	5·05	4·86	4·76	4·65	4·55	4·44	4·34	4·23
12	11·75	8·51	7·23	6·52	6·07	5·76	5·52	5·35	5·20	5·09	4·91	4·72	4·53	4·43	4·33	4·23	4·12	4·01	3·90
13	11·37	8·19	6·93	6·23	5·79	5·48	5·25	5·08	4·94	4·82	4·64	4·46	4·27	4·17	4·07	3·97	3·87	3·76	3·65
14	11·06	7·92	6·68	6·00	5·56	5·26	5·03	4·86	4·72	4·60	4·43	4·25	4·06	3·96	3·86	3·76	3·66	3·55	3·44
15	10·80	7·70	6·48	5·80	5·37	5·07	4·85	4·67	4·54	4·42	4·25	4·07	3·88	3·79	3·69	3·58	3·48	3·37	3·26
16	10·58	7·51	6·30	5·64	5·21	4·91	4·69	4·52	4·38	4·27	4·10	3·92	3·73	3·64	3·54	3·44	3·33	3·22	3·11
17	10·38	7·35	6·16	5·50	5·07	4·78	4·56	4·39	4·25	4·14	3·97	3·79	3·61	3·51	3·41	3·31	3·21	3·10	2·98
18	10·22	7·21	6·03	5·37	4·96	4·66	4·44	4·28	4·14	4·03	3·86	3·68	3·50	3·40	3·30	3·20	3·10	2·99	2·87
19	10·07	7·09	5·92	5·27	4·85	4·56	4·34	4·18	4·04	3·93	3·76	3·59	3·40	3·31	3·21	3·11	3·00	2·89	2·78
20	9·94	6·99	5·82	5·17	4·76	4·47	4·26	4·09	3·96	3·85	3·68	3·50	3·32	3·22	3·12	3·02	2·92	2·81	2·69
21	9·83	6·89	5·73	5·09	4·68	4·39	4·18	4·01	3·88	3·77	3·60	3·43	3·24	3·15	3·05	2·95	2·84	2·73	2·61
22	9·73	6·81	5·65	5·02	4·61	4·32	4·11	3·94	3·81	3·70	3·54	3·36	3·18	3·08	2·98	2·88	2·77	2·66	2·55
23	9·63	6·73	5·58	4·95	4·54	4·26	4·05	3·88	3·75	3·64	3·47	3·30	3·12	3·02	2·92	2·82	2·71	2·60	2·48
24	9·55	6·66	5·52	4·89	4·49	4·20	3·99	3·83	3·69	3·59	3·42	3·25	3·06	2·97	2·87	2·77	2·66	2·55	2·43
25	9·48	6·60	5·46	4·84	4·43	4·15	3·94	3·78	3·64	3·54	3·37	3·20	3·01	2·92	2·82	2·72	2·61	2·50	2·38
26	9·41	6·54	5·41	4·79	4·38	4·10	3·89	3·73	3·60	3·49	3·33	3·15	2·97	2·87	2·77	2·67	2·56	2·45	2·33
27	9·34	6·49	5·36	4·74	4·34	4·06	3·85	3·69	3·56	3·45	3·28	3·11	2·93	2·83	2·73	2·63	2·52	2·41	2·29
28	9·28	6·44	5·32	4·70	4·30	4·02	3·81	3·65	3·52	3·41	3·25	3·07	2·89	2·79	2·69	2·59	2·48	2·37	2·25
29	9·23	6·40	5·28	4·66	4·26	3·98	3·77	3·61	3·48	3·38	3·21	3·04	2·86	2·76	2·66	2·56	2·45	2·33	2·21
30	9·18	6·35	5·24	4·62	4·23	3·95	3·74	3·58	3·45	3·34	3·18	3·01	2·82	2·73	2·63	2·52	2·42	2·30	2·18
40	8·83	6·07	4·98	4·37	3·99	3·71	3·51	3·35	3·22	3·12	2·95	2·78	2·60	2·50	2·40	2·30	2·18	2·06	1·93
60	8·49	5·79	4·73	4·14	3·76	3·49	3·29	3·13	3·01	2·90	2·74	2·57	2·39	2·29	2·19	2·08	1·96	1·83	1·69
120	8·18	5·54	4·50	3·92	3·55	3·28	3·09	2·93	2·81	2·71	2·54	2·37	2·19	2·09	1·98	1·87	1·75	1·61	1·43
∞	7·88	5·30	4·28	3·72	3·35	3·09	2·90	2·74	2·62	2·52	2·36	2·19	2·00	1·90	1·79	1·67	1·53	1·36	1·00

$F=\frac{s_1^2}{s_2^2}=\frac{S_1}{\nu_1}\Big/\frac{S_2}{\nu_2}$, where $s_1^2=S_1/\nu_1$ and $s_2^2=S_2/\nu_2$ are independent mean squares estimating a common variance σ^2 and based on ν_1 and ν_2 degrees of freedom, respectively

Percentage Points of the F-distribution (Variance Ratio) (continued)

Upper 0·1 % points

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4053*	5000*	5404*	5625*	5764*	5859*	5929*	5981*	6023*	6056*	6107*	6158*	6209*	6235*	6261*	6287*	6313*	6340*	6366*
2	998·5	999·0	999·2	999·2	999·3	999·3	999·4	999·4	999·4	999·4	999·4	999·4	999·4	999·5	999·5	999·5	999·5	999·5	999·5
3	167·0	148·5	141·1	137·1	134·6	132·8	131·6	130·6	129·9	129·2	128·3	127·4	126·4	125·9	125·4	125·0	124·5	124·0	123·5
4	74·14	61·25	56·18	53·44	51·71	50·53	49·66	49·00	48·47	48·05	47·41	46·76	46·10	45·77	45·43	45·09	44·75	44·40	44·05
5	47·18	37·12	33·20	31·09	29·75	28·84	28·16	27·64	27·24	26·92	26·42	25·91	25·39	25·14	24·87	24·60	24·33	24·06	23·79
6	35·51	27·00	23·70	21·92	20·81	20·03	19·46	19·03	18·69	18·41	17·99	17·56	17·12	16·89	16·67	16·44	16·21	15·99	15·75
7	29·25	21·69	18·77	17·19	16·21	15·52	15·02	14·63	14·33	14·08	13·71	13·32	12·93	12·73	12·53	12·33	12·12	11·91	11·70
8	25·42	18·49	15·83	14·39	13·49	12·86	12·40	12·04	11·77	11·54	11·19	10·84	10·48	10·30	10·11	9·92	9·73	9·53	9·33
9	22·86	16·39	13·90	12·56	11·71	11·13	10·70	10·37	10·11	9·89	9·57	9·24	8·90	8·72	8·55	8·37	8·19	8·00	7·81
10	21·04	14·91	12·55	11·28	10·48	9·92	9·52	9·20	8·96	8·75	8·45	8·13	7·80	7·64	7·47	7·30	7·12	6·94	6·76
11	19·69	13·81	11·56	10·35	9·58	9·05	8·66	8·35	8·12	7·92	7·63	7·32	7·01	6·85	6·68	6·52	6·35	6·17	6·00
12	18·64	12·97	10·80	9·63	8·89	8·38	8·00	7·71	7·48	7·29	7·00	6·71	6·40	6·25	6·09	5·93	5·76	5·59	5·42
13	17·81	12·31	10·21	9·07	8·35	7·86	7·49	7·21	6·98	6·80	6·52	6·23	5·93	5·78	5·63	5·47	5·30	5·14	4·97
14	17·14	11·78	9·73	8·62	7·92	7·43	7·08	6·80	6·58	6·40	6·13	5·85	5·56	5·41	5·25	5·10	4·94	4·77	4·60
15	16·59	11·34	9·34	8·25	7·57	7·09	6·74	6·47	6·26	6·08	5·81	5·54	5·25	5·10	4·95	4·80	4·64	4·47	4·31
16	16·12	10·97	9·00	7·94	7·27	6·81	6·46	6·19	5·98	5·81	5·55	5·27	4·99	4·85	4·70	4·54	4·39	4·23	4·06
17	15·72	10·66	8·73	7·68	7·02	6·56	6·22	5·96	5·75	5·58	5·32	5·05	4·78	4·63	4·48	4·33	4·18	4·02	3·85
18	15·38	10·39	8·49	7·46	6·81	6·35	6·02	5·76	5·56	5·39	5·13	4·87	4·59	4·45	4·30	4·15	4·00	3·84	3·67
19	15·08	10·16	8·28	7·26	6·62	6·18	5·85	5·59	5·39	5·22	4·97	4·70	4·43	4·29	4·14	3·99	3·84	3·68	3·51
20	14·82	9·95	8·10	7·10	6·46	6·02	5·69	5·44	5·24	5·08	4·82	4·56	4·29	4·15	4·00	3·86	3·70	3·54	3·38
21	14·59	9·77	7·94	6·95	6·32	5·88	5·56	5·31	5·11	4·95	4·70	4·44	4·17	4·03	3·88	3·74	3·58	3·42	3·26
22	14·38	9·61	7·80	6·81	6·19	5·76	5·44	5·19	4·99	4·83	4·58	4·33	4·06	3·92	3·78	3·63	3·48	3·32	3·15
23	14·19	9·47	7·67	6·69	6·08	5·65	5·33	5·09	4·89	4·73	4·48	4·23	3·96	3·82	3·68	3·53	3·38	3·22	3·05
24	14·03	9·34	7·55	6·59	5·98	5·55	5·23	4·99	4·80	4·64	4·39	4·14	3·87	3·74	3·59	3·45	3·29	3·14	2·97
25	13·88	9·22	7·45	6·49	5·88	5·46	5·15	4·91	4·71	4·56	4·31	4·06	3·79	3·66	3·52	3·37	3·22	3·06	2·89
26	13·74	9·12	7·36	6·41	5·80	5·38	5·07	4·83	4·64	4·48	4·24	3·99	3·72	3·59	3·44	3·30	3·15	2·99	2·82
27	13·61	9·02	7·27	6·33	5·73	5·31	5·00	4·76	4·57	4·41	4·17	3·92	3·66	3·52	3·38	3·23	3·08	2·92	2·75
28	13·50	8·93	7·19	6·25	5·66	5·24	4·93	4·69	4·50	4·35	4·11	3·86	3·60	3·46	3·32	3·18	3·02	2·86	2·69
29	13·39	8·85	7·12	6·19	5·59	5·18	4·87	4·64	4·45	4·29	4·05	3·80	3·54	3·41	3·27	3·12	2·97	2·81	2·64
30	13·29	8·77	7·05	6·12	5·53	5·12	4·82	4·58	4·39	4·24	4·00	3·75	3·49	3·36	3·22	3·07	2·92	2·76	2·59
40	12·61	8·25	6·60	5·70	5·13	4·73	4·44	4·21	4·02	3·87	3·64	3·40	3·15	3·01	2·87	2·73	2·57	2·41	2·23
60	11·97	7·76	6·17	5·31	4·76	4·37	4·09	3·87	3·69	3·54	3·31	3·08	2·83	2·69	2·55	2·41	2·25	2·08	1·89
120	11·38	7·32	5·79	4·95	4·42	4·04	3·77	3·55	3·38	3·24	3·02	2·78	2·53	2·40	2·26	2·11	1·95	1·76	1·54
∞	10·83	6·91	5·42	4·62	4·10	3·74	3·47	3·27	3·10	2·96	2·74	2·51	2·27	2·13	1·99	1·84	1·66	1·45	1·00

* Multiply these entries by 100.

This 0·1 % table is based on the following sources: Colcord & Deming (1935); Fisher & Yates (1953, Table V) used with the permission of the authors and of Messrs Oliver and Boyd; Norton (1952).

Percentage Points of the t-distribution

ν	Q = 0·4 2Q = 0·8	0·25 0·5	0·05 0·1	0·025 0·05	0·005 0·01	0·0025 0·005	0·0005 0·001
1	0·325	1·000	6·314	12·706	63·657	127·32	636·62
2	·289	0·816	2·920	4·303	9·925	14·089	31·598
3	·277	·765	2·353	3·182	5·841	7·453	12·924
4	·271	·741	2·132	2·776	4·604	5·598	8·610
5	0·267	0·727	2·015	2·571	4·032	4·773	6·869
6	·265	·718	1·943	2·447	3·707	4·317	5·959
7	·263	·711	1·895	2·365	3·499	4·029	5·408
8	·262	·706	1·860	2·306	3·355	3·833	5·041
9	·261	·703	1·833	2·262	3·250	3·690	4·781
10	0·260	0·700	1·812	2·228	3·169	3·581	4·587
11	·260	·697	1·796	2·201	3·106	3·497	4·437
12	·259	·695	1·782	2·179	3·055	3·428	4·318
13	·259	·694	1·771	2·160	3·012	3·372	4·221
14	·258	·692	1·761	2·145	2·977	3·326	4·140
15	0·258	0·691	1·753	2·131	2·947	3·286	4·073
16	·258	·690	1·746	2·120	2·921	3·252	4·015
17	·257	·689	1·740	2·110	2·898	3·222	3·965
18	·257	·688	1·734	2·101	2·878	3·197	3·922
19	·257	·688	1·729	2·093	2·861	3·174	3·883
20	0·257	0·687	1·725	2·086	2·845	3·153	3·850
21	·257	·686	1·721	2·080	2·831	3·135	3·819
22	·256	·686	1·717	2·074	2·819	3·119	3·792
23	·256	·685	1·714	2·069	2·807	3·104	3·767
24	·256	·685	1·711	2·064	2·797	3·091	3·745
25	0·256	0·684	1·708	2·060	2·787	3·078	3·725
26	·256	·684	1·706	2·056	2·779	3·067	3·707
27	·256	·684	1·703	2·052	2·771	3·057	3·690
28	·256	·683	1·701	2·048	2·763	3·047	3·674
29	·256	·683	1·699	2·045	2·756	3·038	3·659
30	0·256	0·683	1·697	2·042	2·750	3·030	3·646
40	·255	·681	1·684	2·021	2·704	2·971	3·551
60	·254	·679	1·671	2·000	2·660	2·915	3·460
120	·254	·677	1·658	1·980	2·617	2·860	3·373
∞	·253	·674	1·645	1·960	2·576	2·807	3·291

$Q = 1 - P(t|\nu)$ is the upper-tail area of the distribution for ν degrees of freedom, appropriate for use in a single-tail test. For a two-tail test, $2Q$ must be used.

Percentage Points of the X^2-Distribution

ν \ Q	0·995	0·990	0·975	0·950	0·900	0·750	0·500
1	392704.10^{-10}	157088.10^{-9}	982069.10^{-9}	393214.10^{-8}	0·0157908	0·1015308	0·454936
2	0·0100251	0·0201007	0·0506356	0·102587	0·210721	0·575364	1·38629
3	0·0717218	0·114832	0·215795	0·351846	0·584374	1·212534	2·36597
4	0·206989	0·297109	0·484419	0·710723	1·063623	1·92256	3·35669
5	0·411742	0·554298	0·831212	1·145476	1·61031	2·67460	4·35146
6	0·675727	0·872090	1·23734	1·63538	2·20413	3·45460	5·34812
7	0·989256	1·239043	1·68987	2·16735	2·83311	4·25485	6·34581
8	1·34441	1·64650	2·17973	2·73264	3·48954	5·07064	7·34412
9	1·73493	2·08790	2·70039	3·32511	4·16816	5·89883	8·34283
10	2·15586	2·55821	3·24697	3·94030	4·86518	6·73720	9·34182
11	2·60322	3·05348	3·81575	4·57481	5·57778	7·58414	10·3410
12	3·07382	3·57057	4·40379	5·22603	6·30380	8·43842	11·3403
13	3·56503	4·10692	5·00875	5·89186	7·04150	9·29907	12·3398
14	4·07467	4·66043	5·62873	6·57063	7·78953	10·1653	13·3393
15	4·60092	5·22935	6·26214	7·26094	8·54676	11·0365	14·3389
16	5·14221	5·81221	6·90766	7·96165	9·31224	11·9122	15·3385
17	5·69722	6·40776	7·56419	8·67176	10·0852	12·7919	16·3382
18	6·26480	7·01491	8·23075	9·39046	10·8649	13·6753	17·3379
19	6·84397	7·63273	8·90652	10·1170	11·6509	14·5620	18·3377
20	7·43384	8·26040	9·59078	10·8508	12·4426	15·4518	19·3374
21	8·03365	8·89720	10·28293	11·5913	13·2396	16·3444	20·3372
22	8·64272	9·54249	10·9823	12·3380	14·0415	17·2396	21·3370
23	9·26043	10·19567	11·6886	13·0905	14·8480	18·1373	22·3369
24	9·88623	10·8564	12·4012	13·8484	15·6587	19·0373	23·3367
25	10·5197	11·5240	13·1197	14·6114	16·4734	19·9393	24·3366
26	11·1602	12·1981	13·8439	15·3792	17·2919	20·8434	25·3365
27	11·8076	12·8785	14·5734	16·1514	18·1139	21·7494	26·3363
28	12·4613	13·5647	15·3079	16·9279	18·9392	22·6572	27·3362
29	13·1211	14·2565	16·0471	17·7084	19·7677	23·5666	28·3361
30	13·7867	14·9535	16·7908	18·4927	20·5992	24·4776	29·3360
40	20·7065	22·1643	24·4330	26·5093	29·0505	33·6603	39·3353
50	27·9907	29·7067	32·3574	34·7643	37·6886	42·9421	49·3349
60	35·5345	37·4849	40·4817	43·1880	46·4589	52·2938	59·3347
70	43·2752	45·4417	48·7576	51·7393	55·3289	61·6983	69·3345
80	51·1719	53·5401	57·1532	60·3915	64·2778	71·1445	79·3343
90	59·1963	61·7541	65·6466	69·1260	73·2911	80·6247	89·3342
100	67·3276	70·0649	74·2219	77·9295	82·3581	90·1332	99·3341
X	−2·5758	−2·3263	−1·9600	−1·6449	−1·2816	−0·6745	0·0000

$$Q = Q(\chi^2 \mid \nu) = 1 - P(\chi^2 \mid \nu) = 2^{-\frac{1}{2}\nu}\{\Gamma(\tfrac{1}{2}\nu)\}^{-1}\int_{\chi^2}^{\infty} e^{-\frac{1}{2}x} x^{\frac{1}{2}\nu-1}\,dx.$$

Percentage Points of the χ^2-Distribution (continued)

ν \ Q	0·250	0·100	0·050	0·025	0·010	0·005	0·001
1	1·32330	2·70554	3·84146	5·02389	6·63490	7·87944	10·828
2	2·77259	4·60517	5·99146	7·37776	9·21034	10·5966	13·816
3	4·10834	6·25139	7·81473	9·34840	11·3449	12·8382	16·266
4	5·38527	7·77944	9·48773	11·1433	13·2767	14·8603	18·467
5	6·62568	9·23636	11·0705	12·8325	15·0863	16·7496	20·515
6	7·84080	10·6446	12·5916	14·4494	16·8119	18·5476	22·458
7	9·03715	12·0170	14·0671	16·0128	18·4753	20·2777	24·322
8	10·2189	13·3616	15·5073	17·5345	20·0902	21·9550	26·125
9	11·3888	14·6837	16·9190	19·0228	21·6660	23·5894	27·877
10	12·5489	15·9872	18·3070	20·4832	23·2093	25·1882	29·588
11	13·7007	17·2750	19·6751	21·9200	24·7250	26·7568	31·264
12	14·8454	18·5493	21·0261	23·3367	26·2170	28·2995	32·909
13	15·9839	19·8119	22·3620	24·7356	27·6882	29·8195	34·528
14	17·1169	21·0641	23·6848	26·1189	29·1412	31·3194	36·123
15	18·2451	22·3071	24·9958	27·4884	30·5779	32·8013	37·697
16	19·3689	23·5418	26·2962	28·8454	31·9999	34·2672	39·252
17	20·4887	24·7690	27·5871	30·1910	33·4087	35·7185	40·790
18	21·6049	25·9894	28·8693	31·5264	34·8053	37·1565	42·312
19	22·7178	27·2036	30·1435	32·8523	36·1909	38·5823	43·820
20	23·8277	28·4120	31·4104	34·1696	37·5662	39·9968	45·315
21	24·9348	29·6151	32·6706	35·4789	38·9322	41·4011	46·797
22	26·0393	30·8133	33·9244	36·7807	40·2894	42·7957	48·268
23	27·1413	32·0069	35·1725	38·0756	41·6384	44·1813	49·728
24	28·2412	33·1962	36·4150	39·3641	42·9798	45·5585	51·179
25	29·3389	34·3816	37·6525	40·6465	44·3141	46·9279	52·618
26	30·4346	35·5632	38·8851	41·9232	45·6417	48·2899	54·052
27	31·5284	36·7412	40·1133	43·1945	46·9629	49·6449	55·476
28	32·6205	37·9159	41·3371	44·4608	48·2782	50·9934	56·892
29	33·7109	39·0875	42·5570	45·7223	49·5879	52·3356	58·301
30	34·7997	40·2560	43·7730	46·9792	50·8922	53·6720	59·703
40	45·6160	51·8051	55·7585	59·3417	63·6907	66·7660	73·402
50	56·3336	63·1671	67·5048	71·4202	76·1539	79·4900	86·661
60	66·9815	74·3970	79·0819	83·2977	88·3794	91·9517	99·607
70	77·5767	85·5270	90·5312	95·0232	100·425	104·215	112·317
80	88·1303	96·5782	101·879	106·629	112·329	116·321	124·839
90	98·6499	107·565	113·145	118·136	124·116	128·299	137·208
100	109·141	118·498	124·342	129·561	135·807	140·169	149·449
X	+0·6745	+1·2816	+1·6449	+1·9600	+2·3263	+2·5758	+3·0902

For $\nu > 100$ take

$$\chi^2 = \nu\left\{1 - \frac{2}{9\nu} + X\sqrt{\frac{2}{9\nu}}\right\}^3 \quad \text{or} \quad \chi^2 = \tfrac{1}{2}\{X + \sqrt{(2\nu - 1)}\}^2,$$

according to the degree of accuracy required. X is the standardized normal deviate corresponding to $P = 1 - Q$, and is shown in the bottom line of the table.

HEWLETT PACKARD

Part No. 09845-15171
Rev. C

Printed in U.S.A.
May, 1981