## *Step-by-Step Solutions for Your HP-28S or HP-28C Calculator*

*Probability and Statistics* contains a variety of examples and solutions to show how you can easily solve your technical problems.

- **Statistics Matrix Setup**
  Initialization and Data Entry • Data Removal • Data Column Extraction • Data Sets With More Than Two Variables • Grouped Data Matrix Transformation
- **Basic Statistics for Multiple Variables**
  Sums and Means • Standard Deviation, Variance, and Covariance • Computing Sample and Population Statistics • Correlation Coefficient and Coefficient of Variation • Sums of Products • Normalized Data • Delta Percent on Paired Data • Moments, Skewness, and Kurtosis
- **Regression**
  Curve Fitting • Multiple Regressions on the Same Data • Multiple Linear Regression • Polynomial Regression
- **Test Statistics and Confidence Intervals**
  Paired t Statistic • t Statistic for Two Means • Chi-Square Statistic

HEWLETT PACKARD

**Reorder Number**
**00028-90104**

00028-90109 English
Printed in Canada 11/87

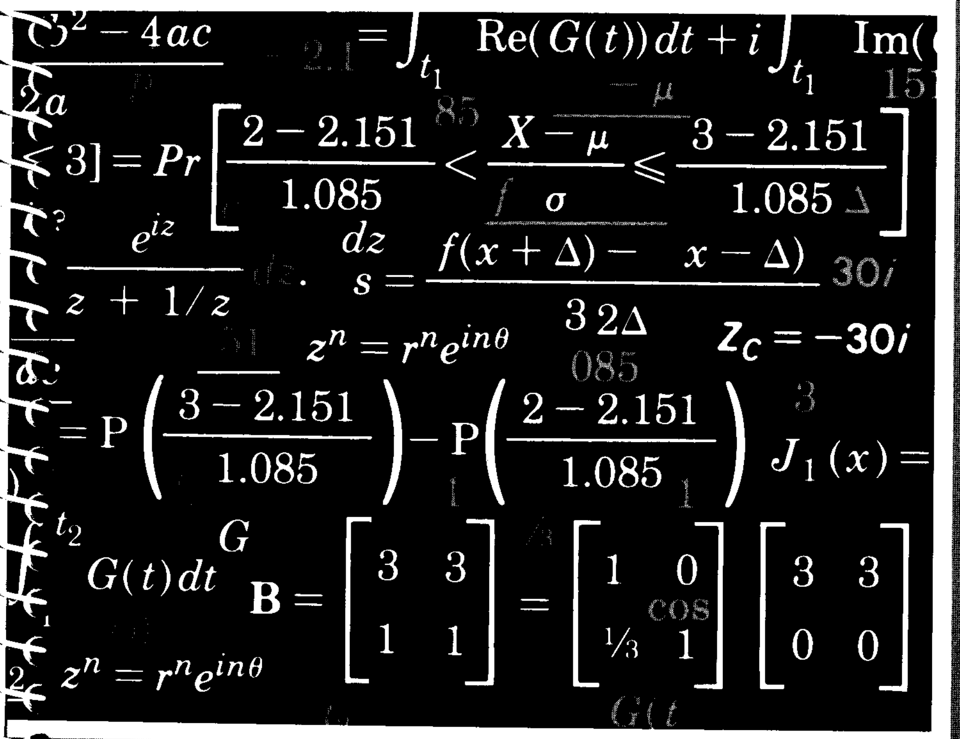
0 88698 00024 3

---

# HEWLETT·PACKARD

Step-by-Step Solutions
For Your HP Calculator

# Probability and Statistics



HP-28S
HP-28C

HEWLETT PACKARD

# Help Us Help You!

Please take a moment to complete this postage-paid card, tear it out and put it in the mail. Your responses and comments will help us better understand your needs and will provide you with the best procedures to solve your problems. Thank you!

## HELP US HELP YOU!

Book: **Probability and Statistics** Date acquired: ____________

Name ____________

Street ____________

City, State, Zip ____________

Phone (______) ____________ Business _____ or Home _____

1. What calculator will you use this book with?
   004 ☐ **HP-28S** 005 ☐ **HP-28C** 006 ☐ **Other** ____________
2. How many other HP solution books have you bought for this calculator? _____
3. What is your OCCUPATION?
   101 ☐ **Student** 103 ☐ **Professional** 109 ☐ **Other** ____________
4. Where did you purchase this book?
   403 ☐ **Bookstore** 404 ☐ **Discount or Catalog Store**
   407 ☐ **Mail Order** 410 ☐ **HP Direct** 411 ☐ **Other** ____________
5. How did you first hear about this book?
   501 ☐ **HP Owner** 503 ☐ **Advertising** 506 ☐ **Salesperson** 507 ☐ **Brochure**
   508 ☐ **Other** ____________
6. To what degree did this book influence your calculator purchase decision?
   601 ☐ **Major Influence** 602 ☐ **Minor Influence** 603 ☐ **No Influence**
7. How well does this book cover the material you expected?
   701 ☐ **Good** 702 ☐ **Moderate** 703 ☐ **Low**
8. What level of knowledge is required to make use of the topics in this book?
   801 ☐ **High** 802 ☐ **Medium** 803 ☐ **Low**
9. How clearly was the material in this book presented?
   901 ☐ **Good** 902 ☐ **Moderate** 903 ☐ **Low**
10. How would you rate the value of this book for your money?
   111 ☐ **High** 112 ☐ **Medium** 113 ☐ **Low**

**Comments:** (Please comment on improvements and additional applications or subjects you would like HP to cover in this or another solution book.) ____________

____________

____________

____________

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY CARD
FIRST CLASS PERMIT NO. 40 CORVALLIS, OR

POSTAGE WILL BE PAID BY ADDRESSEE

Hewlett-Packard Company
ATTN: Calculator Market Research
1000 N.E. Circle Blvd.
Corvallis, OR 97330, U.S.A.

---

# Probability and Statistics

## Step-by-Step Solutions for Your HP-28S or HP-28C Calculator

HEWLETT
PACKARD

Edition 1 November 1987
Reorder Number 00028-90104

## Notice

**Corvallis Division**
**1000 N.E. Circle Blvd.**
**Corvallis, OR 97330, U.S.A.**

## Printing History

**Edition 1** November 1987 Mfg No. 00028-90109

# Welcome...

... to the HP-28S and HP-28C Step-by-Step Solution Books. These books are designed to help you get the most from your HP-28S or HP-28C calculator.

This book, *Probability and Statistics*, provides examples and techniques for solving problems on your calculator. A variety of statistical matrix manipulations and statistical function computations are designed to familiarize you with, and build upon, the statistical capabilities built into your calculator.

Before you try the examples in this book, you should be familiar with certain concepts from the owner's documentation:

- The basics of your calculator: how to move from menu to menu, how to exit graphics and edit modes, and how to use the menu to assign values to, and solve for, user variables.
- Entering numbers, programs, and algebraic expressions into the calculator.

Please review the section "How To Use This Book." It contains important information on the examples in this book.

For more information about the topics in the *Probability and Statistics* book, refer to a basic textbook on the subject. Many references are available in university libraries and in technical and college bookstores. The examples in the book demonstrate approaches to solving certain problems, but they do not cover the many ways to approach solutions to mathematical problems.

# Contents

# How To Use This Book

Please take a moment to familiarize yourself with the formats used in this book.

**Keys and Menu Selection:** A box represents a key on the calculator keyboard.

ENTER
1/x
STO

ARRAY
PLOT
ALGEBRA

In many cases, a box represents a shifted key on the calculator. In the example problems, the shift key is NOT explicitly shown. (For example, ARRAY requires you to press the shift key, followed by the ARRAY key, found above the "A" on the left keyboard.)

The "inverse" highlight represents a menu label:

| **Key:** | **Description:** |
|---|---|
| DRAW | Found in the PLOT menu. |
| ISOL | Found in the SOLV menu. |
| ABCD | A user-created name. If you created a variable by this name, it could be found in either the USER menu or the SOLVR menu. If you created a program by this name, it would be found in the USER menu. |

Menus typically include more menu labels than can be displayed above the six redefinable menu keys. Press NEXT and PREV to roll through the menu options. For simplicity, NEXT and PREV are NOT shown in the examples.

Solving for a user variable within ≣ SOLVR ≣ is initiated by the shift key, followed by the appropriate user-defined menu key:

□ ≣ ABCD ≣.

The keys above indicate the shift key, followed by the user-defined key labeled "ABCD". Pressing these keys initiates the Solver function to seek a solution for "ABCD" in a specified equation.

The symbol <> indicates the cursor-menu key.

**Interactive Plots and the Graphics Cursor:** Coordinate values you obtain from plots using the INS and DEL digitizing keys may differ from those shown, due to small differences in the positions of the graphics cursor. The values you obtain should be satisfactory for the Solver root-finding that follows.

**Display Formats and Numeric Input:** Negative numbers, displayed as

```
-5
-12345.678
[[-1,-2,-3[-4,-5,-6[ ...
```

are created using the CHS key.

```
5 CHS
12345.678 CHS
[[1 CHS,2 CHS, ...
```

The examples in this book typically specify a display format for the number of decimal places. If your display is set such that numeric displays do not match exactly, you can modify your display format with the MODE menu and the ≣ FIX ≣ key within that menu. (For example, MODE 2 ≣ FIX ≣ will set the display to the FIX 2 format.)

**Programming Reminders:** Before you key in the programming examples in this book, familiarize yourself with the locations of programming commands that appear as menu labels. By using the menu labels to enter commands, you can speed keying in programs and avoid errors that might arise from extra spaces appearing in the programs. Remember, the calculator recognizes commands that are set off by spaces. Therefore, the arrow ( → ) in the command R→C (the real to complex conversion function) is interpreted differently than the arrow in the command → C (create the local variable "C").

The HP-28S automatically inserts spaces around each operator as you key it in. Therefore, using the R, →, and C keys to enter the R→C command will result in the expression R → C , and, ultimately, in an error in your program. As you key in programs on the HP-28S, take particular care to avoid spaces inside commands, especially in commands that include an →.

The HP-28C does not automatically insert spaces around operators or commands as they are keyed in.

**A Note About the Displays Used in This Book:** The menus and screens that appear in this book show the HP-28S display. Most of the HP-28C and HP-28S screens are identical, but there are differences in the MODE menu and ≣ SOLVR ≣ screen that HP-28C users should be aware of.

For example, the first screen below illustrates the HP-28C MODE menu, and the second screen illustrates the same menu as it appears on the HP-28S.

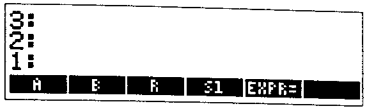
HP-28C MODE display.

HP-28S MODE display.

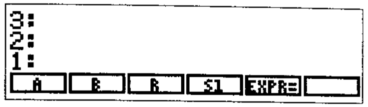
Notice that the HP-28C highlights the entire active menu item, while the HP-28S display includes a small box in the active menu item.

The screens shown below illustrate the HP-28C and HP-28S versions of the SOLVR menu.

HP-28C SOLVR display.



HP-28S SOLVR display.



Both of these screens include the Solver variables A, B, R, S1, and EXPR=. The HP-28C displays Solver variables in gray on a black background. The HP-28S prints Solver variables in black on a gray background.

**User Menus:** A PURGE command follows many of the examples in this book. If you do not purge all of the programs and variables after working each example, or if your USER menu contains your own user-defined variables or programs, the USER menu on your calculator may differ from the displays shown in this book. Do not be concerned if the variables and programs appear in a slightly different order on your USER menu; this will not affect the calculator's performance.

# Statistics Matrix Setup

This section describes the structure of the statistical matrices used in the remainder of this book, and provides a number of techniques for manipulating the data within the matrix. An approach to managing grouped data is also described.

# Initialization and Data Entry

The statistical calculations throughout this book generally operate on single or paired columns of data collected in the variable ΣDAT. This statistical matrix can, however, hold additional data vectors, providing for multiple pairing and analysis.

**Ungrouped Data Matrix:** The statistical matrix ΣDAT for ungrouped data has the form shown below.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

The matrix shown above has $n$ sets (vectors) of statistical data, each containing $m$ data points.

**Grouped Data Matrix:** The approach used to manage grouped data in this book is to collect data using the same functions as for ungrouped data, including data entry and removal and data pair selection. However, once the grouped data has been entered, the data matrix is stored in another matrix variable, and the data is expanded into ΣDAT as if it were ungrouped data. This approach has a disadvantage in terms of memory consumption since the data is effectively retained twice in the machine. However, it greatly simplifies the steps and programs to compute basic and advanced statistics on the data since many powerful functions for ungrouped data are built into the HP-28S and HP-28C.

Thus, the grouped data matrix shown below is transformed to the ungrouped form shown earlier prior to calculating statistics for the data.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} & g_1 \\ x_{21} & x_{22} & \cdots & x_{2m} & g_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} & g_n \end{bmatrix}$$

Data vector one in the matrix above occurs $g_1$ times, data vector two occurs $g_2$ times, and so on.

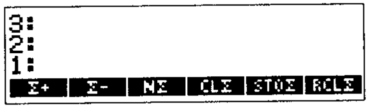
## Initialization and Data Entry Examples

At the beginning of each new problem, the current statistical matrix is cleared by pressing CLΣ. A statistical matrix may also be saved for later use by recalling and assigning it to another matrix variable.

**Example:** Clear the stack and the current statistical matrix, and then enter the following ungrouped statistical data.

$$\begin{bmatrix} x_i & y_i \\ -- & -- \\ 26 & 92 \\ 30 & 85 \\ 44 & 78 \\ 50 & 81 \\ 62 & 54 \\ 68 & 51 \\ 74 & 40 \end{bmatrix}$$

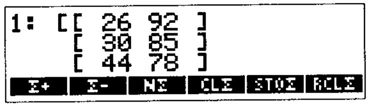
CLEAR STAT CLΣ
[26,92 Σ+
[30,85 Σ+
[44,78 Σ+
[50,81 Σ+
[62,54 Σ+
[68,51 Σ+
[74,40 Σ+



```
3:
2:
1:
```

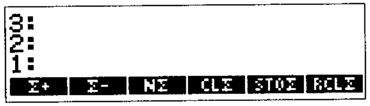
Now recall the statistical matrix and copy it to another variable named *UNGR*.

RCLΣ



```
1: [[ 26 92 ]
    [ 30 85 ]
    [ 44 78 ]
```

'UNGR STO

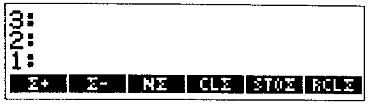


```
3:
2:
1:
```

**Example:** Clear the current statistical matrix and enter the following set of grouped data.

$$\begin{bmatrix} x_i & y_i & g_i \\ \text{--} & \text{--} & \text{--} \\ 4.8 & 15.1 & 1 \\ 5.2 & 11.5 & 3 \\ 3.8 & 14.3 & 1 \\ 4.4 & 13.6 & 6 \\ 4.1 & 12.8 & 2 \end{bmatrix}$$

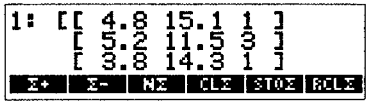
CLΣ
`[4.8,15.1,1` Σ+
`[5.2,11.5,3` Σ+
`[3.8,14.3,1` Σ+
`[4.4,13.6,6` Σ+
`[4.1,12.8,2` Σ+



```
3:
2:
1:
```

Recall this matrix of grouped data, copy to the variable *GRUP*, and redisplay the matrix.

RCLΣ `'GRUP` STO
RCLΣ



```
1: [[ 4.8 15.1 1 ]
    [ 5.2 11.5 3 ]
    [ 3.8 14.3 1 ]
```

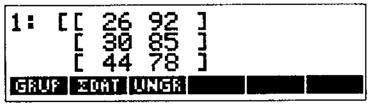
## Data Removal

The last data vector in the statistical matrix is easily removed by using Σ-. Removal of a specified vector other than the last entry is accomplished with the program on page 16.

**Last Data Vector Removal:** Remove the last data vector from the ungrouped data matrix and display and view the matrix contents.

Clear the stack and recall the ungrouped data from the variable *UNGR*.

CLEAR USER UNGR



```
1: [[ 26 92 ]
    [ 30 85 ]
    [ 44 78 ]
```

Make this matrix the current statistical matrix.

STAT STOΣ

```
3:
2:
1:
```

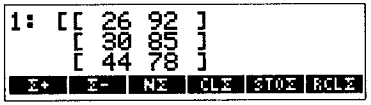
Now remove the last data vector.

Σ-

```
3:
2:
1:                 [ 74 40 ]
```

Recall the matrix to the stack and examine the contents.

RCLΣ
VIEW↓
VIEW↓
VIEW↓
ATTN



```
1: [[ 26 92 ]
    [ 30 85 ]
    [ 44 78 ]
```

Note that there are now six, rather than seven, pairs of data in the statistical matrix. *UNGR* still contains the original data set.

**Arbitrary Data Vector Removal:** Key in the program below for removing a specified data vector from the current statistical matrix.

| **Program:** | **Comments:** |
|---|---|
| `« NΣ SWAP - → n` | Compute number of rows below row to be discarded. |
| `« 1 n 1 + START Σ-`<br>`NEXT` | Stack rows from the bottom through the discard row. |
| `DROP` | Drop the discard row. |
| `IF n 0 ≠ THEN 1 n`<br>`START Σ+ NEXT END »`<br>`»` | Put the rows below the discard row back into the statistical matrix. |

ENTER 'DELI STO

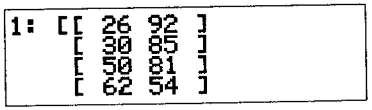
**Example:** Remove the third data vector from the current statistical matrix.

CLEAR
3 USER DELI

```
3:
2:
1:
```

Display the matrix.

STAT RCLΣ <>



```
1: [[ 26 92 ]
    [ 30 85 ]
    [ 50 81 ]
    [ 62 54 ]
```

## Data Column Extraction

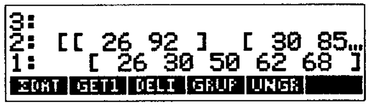
The program *GET1* retrieves a column of data from the current statistical matrix. The desired column number is passed to the program in level one of the stack.

| **Program:** | **Comments:** |
|---|---|
| `« RCLΣ DUP TRN STOΣ` | Save the original data and transpose to drop unwanted columns in row form. |
| `SWAP 1 + NΣ DUP2 IF`<br>`≤ THEN START Σ- DROP`<br>`NEXT ELSE DROP2 END` | Drop all rows (columns) beyond specified one. |
| `Σ- SWAP STOΣ »` | Get specified column; restore data. |

ENTER 'GET1 STO

**Example:** Get a vector containing the elements of the first column of the current statistical matrix.

1 USER GET1



```
3:
2: [[ 26 92 ]  [ 30 85…
1:    [ 26 30 50 62 68 ]
```

## Data Sets With More Than Two Variables

The user variable ΣPAR contains a list of four real numbers. The first two numbers determine the columns of the statistical matrix operated on by the statistical functions of the calculator.

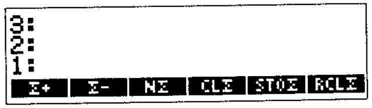
The command COLΣ takes two column numbers from the stack and stores them as the first two objects in the list in the variable ΣPAR.

**Example:** For the multiple-pair data set below, specify columns one and three as the pair for analysis. (This capability will be discussed further in later examples in this book.)

$$\begin{bmatrix} t_i & x_i & y_i \\ \text{--} & \text{--} & \text{--} \\ 1 & 26 & 92 \\ 2 & 30 & 85 \\ 3 & 44 & 78 \\ 4 & 50 & 81 \\ 5 & 62 & 54 \\ 6 & 68 & 51 \\ 7 & 74 & 40 \end{bmatrix}$$

Clear the stack and the current statistical matrix, and enter the data above.

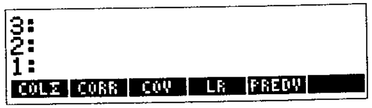
CLEAR STAT CLΣ
[1,26,92 Σ+
[2,30,85 Σ+
[3,44,78 Σ+
[4,50,81 Σ+
[5,62,54 Σ+
[6,68,51 Σ+
[7,74,40 Σ+



```
3:
2:
1:
```

Now specify columns one and three as the pair of data vectors for analysis.

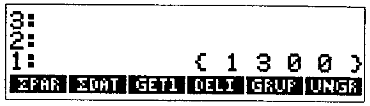
1,3 COLΣ



```
3:
2:
1:
```

Recall the statistical matrix parameters to examine the columns specified.

USER ΣPAR



```
3:
2:
1:                { 1 3 0 0 }
```

Columns one and three are the currently specified data vectors. This process is useful for multiple data vector manipulations or regressions on multiple sets of data with the same base- or time-line.

## Grouped Data Matrix Transformation

The program below transforms grouped data in the current statistical matrix to an ungrouped form in the current statistical matrix.

| **Program:** | **Comments:** |
|---|---|
| `« RCLΣ 'GD' STO RCLΣ` | Recall the grouped data and save it in *GD*. |
| `CLΣ ARRY→ LIST→ DROP`<br>`→ n m` | Transform grouped data to element form and save the dimensions. |
| `« 1 n START` | Loop *n* times. |
| `m ROLLD m 1 - 1`<br>`→LIST →ARRY → ar` | Save $g_i$ in stack; place data in temporary vector for expansion. |
| `« 1 SWAP START ar Σ+`<br>`NEXT »` | Setup and accumulate *ar* $g_i$ times. |
| `NEXT » »` | Repeat outer loop. |

ENTER 'XFRM STO

**Example:** Use the program *XFRM* to transform the grouped data matrix *GRUP*.

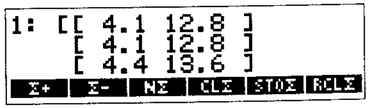
With the program entered and stored, recall the grouped data matrix and make it the current statistical matrix.

CLEAR
USER GRUP
STAT STOΣ

```
3:
2:
1:
```

Now run the program *XFRM* on the grouped data and recall the current statistical matrix to review the data.

USER XFRM
STAT RCLΣ



```
1: [[ 4.1 12.8 ]
    [ 4.1 12.8 ]
    [ 4.4 13.6 ]
```

Use VIEW↓ to scan the matrix. Note how the program builds the transformed, ungrouped matrix from the bottom to the top of the grouped data.

Purge the grouped data matrix *GD*. The original data exists in *GRUP*.

'GD PURGE

# Basic Statistics for Multiple Variables

This chapter provides keystrokes and programs to calculate a variety of basic statistics on the current statistical matrix. These statistics include mean, standard deviation, variance, and covariance on both samples and populations. Techniques for calculating the correlation coefficient, coefficient of variation, sums of products, normalized data, moments, and delta percents on paired statistics are also included.

The current statistical matrix is assumed to be ungrouped data in the calculations that follow. For grouped data, the statistical matrix should be transformed by the program *XFRM* described in the previous chapter.

## Sums and Means

Sums and means for each column of statistics in the current statistical matrix are easily calculated on the HP-28S or HP-28C. The mean and sum are computed from the formulas

$$\text{mean} = \sum_{i=1}^{n} \frac{x_i}{n}$$

and

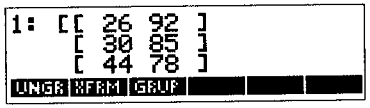
$$\text{sum} = \sum_{i=1}^{n} x_i$$

where $x_i$ is the $i$th coordinate value in a column, and $n$ is the number of data vectors.

**Example:** Compute the sums and means for the ungrouped data stored in the variable *UNGR*.

First, clear the stack and recall the matrix. Remember, your `USER` display may differ from the display shown here.

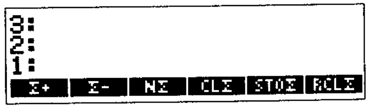
`CLEAR` `USER` `UNGR`



```
1: [[ 26 92 ]
    [ 30 85 ]
    [ 44 78 ]
```

Specify this matrix as the current statistical matrix.

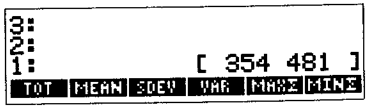
`STAT` `STOΣ`



```
3:
2:
1:
```

Compute the column totals.

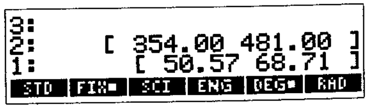
`TOT`



```
3:
2:
1:                  [ 354 481 ]
```

Compute the means for the two columns. Change the display setting to two digits following the decimal point.

`MEAN` `MODE` 2 `FIX`



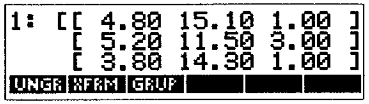
```
3:
2:      [ 354.00 481.00 ]
1:          [ 50.57 68.71 ]
```

**Example: Weighted Mean.** For grouped data, the statistical matrix should be transformed by the program *XFRM* described on page 20.

Compute the weighted mean for the grouped data stored in the variable *GRUP*.

First, clear the stack and recall the matrix.

`STAT` `CLΣ`
`CLEAR` `USER` `GRUP`



```
1: [[ 4.80 15.10 1.00 ]
    [ 5.20 11.50 3.00 ]
    [ 3.80 14.30 1.00 ]
```

Specify this matrix as the current statistical matrix.

`STAT` `STOΣ`

```
3:
2:
1:
```

Transform the matrix and compute the column totals.

`USER` `XFRM`
`STAT` `TOT`

```
3:
2:
1:      [ 58.80 171.10 ]
```

Compute the mean. The weighted mean of the grouped data is simply the mean of the transformed grouped data matrix.

`MEAN`

```
3:
2:      [ 58.80 171.10 ]
1:        [ 4.52 13.16 ]
```

# Standard Deviation, Variance, and Covariance

Both sample and population statistics are readily computed using the built-in functions of the HP-28S or HP-28C. For the population statistics, a short program described in the "STAT" section of the reference manual makes calculation of the population statistics easy.

## Standard Deviation

The standard deviation of the sample and population are given by the following formulas.

### Sample Standard Deviation:

$$s_x = \left( \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 \right)^{\frac{1}{2}}$$

### Population Standard Deviation:

$$\sigma_x = \left( \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 \right)^{\frac{1}{2}}$$

## Variance

The variance of the sample and population are given by the following formulas.

### Sample Variance:

$$s_x{}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

### Population Variance:

$$\sigma_x{}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

For the formulas above, $x_i$ is the $i$th coordinate value in a column, $\overline{x}$ is the mean of the data in this column, and $n$ is the number of data vectors.

## Covariance

The covariance of the sample and population are given by the following formulas.

### Sample Covariance:

$$s_{xy} = \frac{1}{n-1} \left( \sum_{i=1}^{n} (x_{im_1} - \overline{x}_{im_1})(x_{im_2} - \overline{x}_{n_2}) \right)$$

### Population Covariance:

$$\sigma_{xy} = \frac{1}{n} \left( \sum_{i=1}^{n} (x_{im_1} - \overline{x}_{im_1})(x_{im_2} - \overline{x}_{n_2}) \right)$$

In the above formulas, $n$ is the number of data vectors, $x_{im_1}$ is the $i$th coordinate value in column $m_1$, and $\overline{x}_{m_1}$ is the mean of the data in column $m_1$.

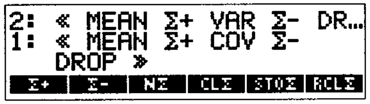
## Computing Sample and Population Statistics

A program of the general form `« MEAN Σ+ fn Σ- DROP »`, where `"fn"` is replaced by the appropriate HP-28S or HP-28C function (`SDEV`, `VAR`, or `COV`), will compute the population statistics for the specified function.

**Example:** Compute the sample and population standard deviation, variance, and covariance for the ungrouped statistical matrix.

First, key in the population statistics programs. (A fast way to key in the second and third programs is to duplicate and EDIT the previous program with the function change.)

```
CLEAR  STAT
« MEAN Σ+ SDEV Σ- DROP
ENTER
« MEAN Σ+ VAR Σ- DROP
ENTER
« MEAN Σ+ COV Σ- DROP
ENTER
```



```
2: « MEAN Σ+ VAR Σ- DR…
1: « MEAN Σ+ COV Σ-
   DROP »
```

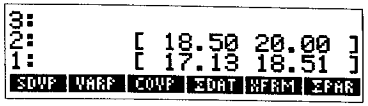
Store the population statistics programs.

```
'COVP  STO
'VARP  STO
'SDVP  STO  USER
```

```
3:
2:
1:
```

Now compute the sample and population standard deviations.

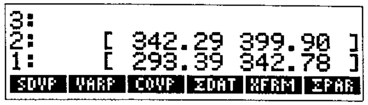
CLEAR UNGR
STAT STOΣ SDEV
USER SDVP



```
3:
2:          [ 18.50 20.00 ]
1:          [ 17.13 18.51 ]
SDVP VARP COVP ΣDAT XFRM ΣPAR
```

Compute the sample and population variances.

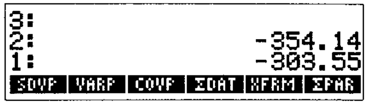
CLEAR
STAT VAR
USER VARP



```
3:
2:     [ 342.29 399.90 ]
1:     [ 293.39 342.78 ]
SDVP VARP COVP ΣDAT XFRM ΣPAR
```

Compute the sample and population covariances. Specify columns one and two in *ΣPAR* if they have been set otherwise.

CLEAR
STAT 1,2 COLΣ
COV
USER COVP



```
3:
2:                 -354.14
1:                 -303.55
SDVP VARP COVP ΣDAT XFRM ΣPAR
```

## Correlation Coefficient and Coefficient of Variation

The correlation coefficient and coefficient of variation are computed by the following formulas.

### Correlation Coefficient:

$$\frac{\sum_{i=1}^{n}(x_{im_1}-\overline{x}_{m_1})(x_{im_2}-\overline{x}_{m_2})}{\left(\sum_{i=1}^{n}(x_{im_1}-\overline{x}_{m_1})^2\sum_{i=1}^{n}(x_{im_2}-\overline{x}_{m_2})^2\right)^{\frac{1}{2}}}$$

### Coefficient of Variation:

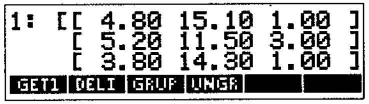
$$V_x=\frac{s_x}{\overline{x}}\cdot 100$$

The terms are defined in the previous problem section.

Compute the statistics above for the *grouped* statistical data *GRUP*.

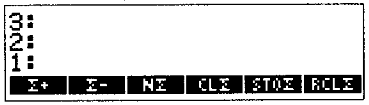
First, recall and transform the grouped data into the current statistical matrix.

CLEAR
USER GRUP



```
1: [[ 4.80 15.10 1.00 ]
    [ 5.20 11.50 3.00 ]
    [ 3.80 14.30 1.00 ]
GET1 DELI GRUP UNGR
```

STAT STOΣ



```
3:
2:
1:
Σ+ Σ- NΣ CLΣ STOΣ RCLΣ
```

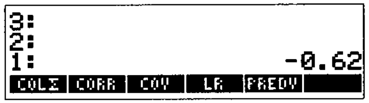
Now transform the grouped data into its ungrouped form.

USER XFRM

```
3:
2:
1:
ΣDAT GD SDVP VARP COVP XFRM
```

Compute the statistics.

STAT CORR

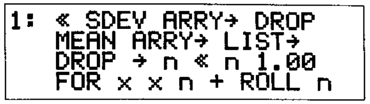


```
3:
2:
1:                  -0.62
COLΣ CORR COV LR PREDV
```

The correlation coefficient is calculated with a built-in function.

Coefficients of variation are calculated by entering the following program.

```
« SDEV ARRY→ DROP
MEAN ARRY→ LIST→
DROP → n « n 1 FOR x
x n + ROLL n 1 +
ROLL ÷ 100 × -1 STEP
n 1 →LIST →ARRY
ENTER  <>
```



```
1: « SDEV ARRY→ DROP
   MEAN ARRY→ LIST→
   DROP → n « n 1.00
   FOR x x n + ROLL n
```

Store the program in the variable *VCO*.

```
'VCO STO
CLEAR
```

```
4:
3:
2:
1:
```

Compute the coefficients of variation.

```
USER  VCO
```

```
3:
2:
1:            [ 9.93 8.42 ]
 VCO  ΣDAT  SDVP  VARP  COVP  NFRM
```

The statistical matrices *GRUP*, *GD*, and *UNGR* are not used in further examples. Purge them and the variables *VCO*, *SDVP*, *VARP*, and *COVP*.

```
{'GD''GRUP''UNGR''VCO''SDVP''VARP''COVP'
PURGE
```

---

## Sums of Products

The HP-28S and HP-28C matrix functions provide an easy method of computing the sums of products of statistical data. A matrix multiplied by its transposed matrix will produce the result shown below.

Let the current statistical matrix be

$$M = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

Then $M^T \cdot M$ is an $m \times m$ matrix of the form

$$\begin{bmatrix} \sum x_{i1}^2 & \sum x_{i1}x_{i2} & \cdots & \sum x_{i1}x_{im} \\ \sum x_{i1}x_{i2} & \sum x_{i2}^2 & \cdots & \sum x_{i2}x_{im} \\ \cdots & \cdots & \cdots & \cdots \\ \sum x_{i1}x_{im} & \sum x_{i2}x_{im} & \cdots & \sum x_{im}^2 \end{bmatrix}$$

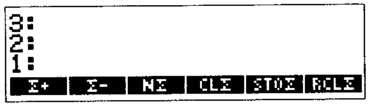
If $M$ represents columns of statistics then each element of $M^T \cdot M$ represents the sum of products of two columns of statistics such that all pairings of columns in $M$ are accounted for.

**Example:** Compute the sums of products on all pairings of the following data.

$$\begin{bmatrix} 10 & 4 & 7 \\ 20 & 5 & 3 \\ 3 & 9 & 2 \\ 5 & 2 & 1 \\ 7 & 4 & 5 \end{bmatrix}$$

Key in the statistical data.

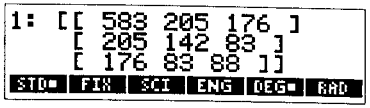
```
CLEAR
STAT  CLΣ
[10,4,7  Σ+
[20,5,3  Σ+
[3,9,2  Σ+
[5,2,1  Σ+
[7,4,5  Σ+
```



```
3:
2:
1:
 Σ+   Σ-   NΣ   CLΣ  STOΣ  RCLΣ
```

Compute the sums of products.

RCLΣ RCLΣ
ARRAY TRN SWAP ×
MODE STD



```
1: [[ 583 205 176 ]
    [ 205 142 83 ]
    [ 176 83 88 ]]
 STD■  FIX  SCI  ENG  DEG■  RAD
```

---

## Normalized Data

A column of data in the current statistical matrix can be normalized by transforming each element as shown below.

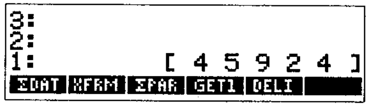
$$x_j{}' = \frac{x_j}{\sum_{i=1}^{n} x_i}$$

$x_j{}'$ is created by dividing the original $x_j$ by the sum of the column data.

**Example:** Compute a normalized vector from the second column of the data of the previous section.

First, use program *GET1* to extract the second column.

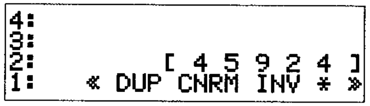
CLEAR
2 USER GET1



```
3:
2:
1:            [ 4 5 9 2 4 ]
 ΣDAT  XFRM  ΣPAR  GET1  DELI
```

Key in the program below. The program divides each element of the vector by the sum of the absolute values of each of the elements.

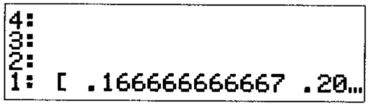
« DUP CNRM INV ×
ENTER <>



```
4:
3:
2:           [ 4 5 9 2 4 ]
1:    « DUP CNRM INV * »
```

Now execute the program.

EVAL



```
4:
3:
2:
1: [ .166666666667 .20…
```

ARRY→ will break the vector into component form for examination. The program can be stored for repeated use if desired.

# Delta Percent on Paired Data

The delta percent of a pair of columnar data can be computed by the program below.

$$\Delta\% = \frac{new - old}{old}$$

*Old* represents the first column of data specified in ΣPAR, and *new* represents the second column of data specified in ΣPAR.

## Delta Percent Program

For this program to work properly, you must first create the variable *ΣPAR*. If *ΣPAR* does not appear in the USER menu, you can create it by the keystrokes `1,2 COLΣ`. If you are working with only two columns, you can simplify the program below by removing the flexibility to specify the columns to be used in the computation; i.e., `ΣPAR 2 GET` can be replaced by the column number desired, and similarly, `ΣPAR 1 GET` can be replaced by the other column number.)

The program below assumes the program *GET1* is already resident in the calculator.

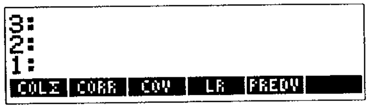
| **Program:** | **Comments:** |
|---|---|
| `« ΣPAR 2 GET GET1`<br>`ARRY→ DROP ΣPAR 1`<br>`GET GET1 ARRY→ DROP` | Get the data from the two columns on the stack. |
| `NΣ 1 FOR x x NΣ +`<br>`ROLL → new` | Roll down and store *new*. |
| `« NΣ ROLL → old` | Roll down and store *old*. |
| `« '(new-old)÷old'`<br>`EVAL » »` | Compute the delta percent. |
| `-1 STEP NΣ 1 2 →LIST`<br>`→ARRY »` | Count down until complete, then structure the data into a column. |

ENTER `'DLTA` STO

**Example:** Compute the delta percent between columns one and three of the data originally entered in the Sums of Products, on page 29.

Select columns one and three.

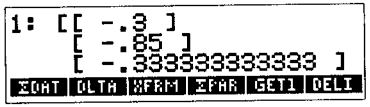
CLEAR
`1,3` STAT COLΣ



```
3:
2:
1:
 COLΣ  CORR  COV   LR   PREDV
```

Now compute the delta percent between the pair of columns.

USER DLTA



```
1: [[ -.3 ]
    [ -.85 ]
    [ -.333333333333 ]
 ΣDAT  DLTA  XFRM  ΣPAR  GET1  DELI
```

Purge variables created in this section.

`{'DLTA''XFRM''DELI''ΣPAR` PURGE

## Moments, Skewness, and Kurtosis

For grouped or ungrouped data, moments are used to describe sets of data, skewness is used to measure the lack of symmetry in a distribution, and kurtosis is the relative peakness or flatness of a distribution.

For a given set of data

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

the moments and moment coefficients are calculated by the following expressions.

**1st Moment:**

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

**2nd Moment:**

$$m_2 = \frac{1}{n}\sum x_i^2 - \overline{x}^2$$

**3rd Moment:**

$$m_3 = \frac{1}{n}\sum x_i^3 - \frac{3}{n}\overline{x}\sum x_i^2 + 2\overline{x}^3$$

**4th Moment:**

$$m_4 = \frac{1}{n}\sum x_i^4 - \frac{4}{n}\overline{x}\sum x_i^3 + \frac{6}{n}\overline{x}^2\sum x_i^2 - 3\overline{x}^4$$

**Moment Coefficient of Skewness:**

$$\gamma_1 = \frac{m_3}{m_2^{3/2}}$$

**Moment Coefficient of Kurtosis:**

$$\gamma_2 = \frac{m_4}{m_2^2}$$

### Calculating Moments, Skewness and Kurtosis

Two programs for computing moments, skewness, and kurtosis are described below. The first program requires specification of a column of data for the current statistical matrix. It calls the second program repeatedly to compute various sums of powers of the columnar data.

The program below assumes the program *GET1* is already resident in the calculator.

| **Program:** | **Comments:** |
|---|---|
| `« GET1 NΣ 1 2 →LIST`<br>`RDM` | Retrieve the specified column and transform it into a column vector. |
| `RCLΣ SWAP STOΣ` | Save the current statistical matrix on the stack; store the column vector. |
| `2 SUMO NΣ ÷ MEAN SQ`<br>`- 'M2' STO` | Compute the second moment. |
| `3 SUMO 2 SUMO MEAN ×`<br>`3 × - NΣ ÷ MEAN 3 ^`<br>`2 × + 'M3' STO` | Compute the third moment. |
| `4 SUMO 3 SUMO 4 × 2`<br>`SUMO 6 × MEAN × -`<br>`MEAN × - NΣ ÷ MEAN 4`<br>`^ 3 × - 'M4' STO` | Compute the fourth moment. |
| `'M3÷M2^1.5' EVAL`<br>`'GM1' STO` | Compute the coefficient of skewness. |
| `'M4÷M2^2' EVAL 'GM2'`<br>`STO` | Compute the coefficient of kurtosis. |
| `STOΣ »` | Restore the complete statistical array. |

```
ENTER 'MOMS STO
```

| **Program:** | **Comments:** |
|---|---|
| `« → p « RCLΣ ARRY→`<br>`DROP` | Store the power for $\sum x^p$ and place the data separately on the stack. |
| `0 1 NΣ START` | Zero for sum accumulation, set up loop count. |
| `SWAP p ^ +` | Compute $x_i^p$ and sum. |
| `NEXT » »` | Complete loop, end program. |

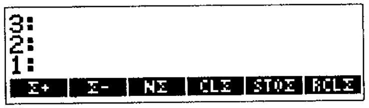
```
ENTER 'SUMO STO
```

**Example:** Compute the first through fourth moments and the coefficients of skewness and kurtosis for the data below.

| $x_i$ | $y_i$ |
|---|---|
| --- | --- |
| 2.1 | 1.1 |
| 3.5 | 3.8 |
| 4.2 | 4.4 |
| 6.5 | 9.7 |
| 4.1 | 3.2 |
| 3.6 | 2.2 |
| 5.3 | 1.6 |
| 3.7 | 5.0 |
| 4.9 | 1.7 |

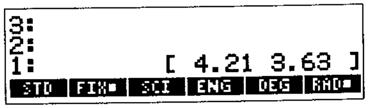
Clear the current statistical matrix and enter the data.

CLEAR STAT CLΣ
`[2.1,1.1` Σ+
`[3.5,3.8` Σ+
`[4.2,4.4` Σ+
`[6.5,9.7` Σ+
`[4.1,3.2` Σ+
`[3.6,2.2` Σ+
`[5.3,1.6` Σ+
`[3.7,5` Σ+
`[4.9,1.7` Σ+


```
3:
2:
1:
 Σ+  | Σ-  | NΣ  | CLΣ | STOΣ | RCLΣ
```

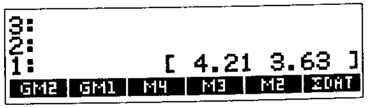
Compute the first moments.

MEAN MODE 2 FIX


```
3:
2:
1:            [ 4.21 3.63 ]
 STD | FIX■ | SCI | ENG | DEG | RAD■
```

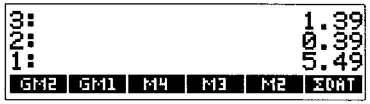
Specify the first column and compute the other $x_i$ moments and coefficients.

1 USER MOMS


```
3:
2:
1:            [ 4.21 3.63 ]
 GM2 | GM1 | M4 | M3 | M2 | ΣDAT
```

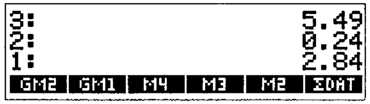
Display the second, third, and fourth moments.

M2 M3 M4


```
3:                      1.39
2:                      0.39
1:                      5.49
 GM2 | GM1 | M4 | M3 | M2 | ΣDAT
```

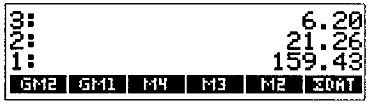
Display the coefficients of skewness and kurtosis.

GM1 GM2

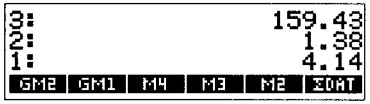

```
3:                      5.49
2:                      0.24
1:                      2.84
 GM2 | GM1 | M4 | M3 | M2 | ΣDAT
```

Repeat the process for the second column of data $y_i$.

CLEAR 2 MOMS
M2 M3 M4


```
3:                      6.20
2:                     21.26
1:                    159.43
 GM2 | GM1 | M4 | M3 | M2 | ΣDAT
```

GM1 GM2


```
3:                    159.43
2:                      1.38
1:                      4.14
 GM2 | GM1 | M4 | M3 | M2 | ΣDAT
```

Purge variables created in this section.

`{'MOMS''SUMO''M2''M3''M4''GM1''GM2` PURGE

# Regression

A variety of regression techniques are performed easily with either the HP-28S or HP-28C. The calculator's built-in matrix manipulation and system solution capabilities, coupled with data and curve plotting, make it a very capable tool for regression analysis.

## Curve Fitting

This problem section describes programs to compute linear, exponential, logarithmic, and power curve fits to a set of data points in the current statistical matrix. Any or all of the curve types may be selected to find a best fit. The data and regression equation may be plotted, and estimates from the regression equation are easily computed with the Solver.

The programs and instructions that follow are designed for flexibility in trying different types of regressions on the same data. If your analysis requirements are for linear regression only, you should use the built-in commands for linear regression, described in the owner's documentation.

For a set of data points $(x_i, y_i)$, the regression equations for four types of curves are shown below.

**Straight Line (Linear Regression):**

$$y = a + bx$$

**Exponential Curve:**

$$y = ae^{bx} \quad \text{where } a > 0$$

**Logarithmic Curve:**

$$y = a + b\ln(x)$$

**Power Curve:**

$$y = ax^{b} \quad \text{where } a > 0$$

The regression coefficients $a$ and $b$ are found by solving the following system of linear equations.

$$\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} A \\ b \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum Y_i X_i \end{bmatrix}$$

where the variables are defined below.

| **Regression** | $A$ | $X_i$ | $Y_i$ |
|---|---|---|---|
| Linear | $a$ | $x_i$ | $y_i$ |
| Exponential | $\ln a$ | $x_i$ | $\ln y_i$ |
| Logarithmic | $a$ | $\ln x_i$ | $y_i$ |
| Power | $\ln a$ | $\ln x_i$ | $\ln y_i$ |

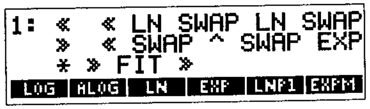
The coefficient of determination is

$$R^2 = \frac{A\sum Y_i + b\sum X_i Y_i - \frac{1}{n}(\sum Y_i)^2}{\sum(Y_i^2) - \frac{1}{n}(\sum Y_i)^2}$$

The programs below apply the least squares method, either to the original data or the transformed data, as described above. For all regression types, the original data is restored after the computation of the regression equation. This allows for multiple regression types to be tried on the same data set.

Key in the four programs below. These programs define the data transformations and the equation-generating transforms for the general curve fitting program. For convenience, change to the [LOGS] menu and use the menu items as you key in the program. Notice that it is not necessary to key the closing delimiter "»" before you press [ENTER].

```
[LOGS]
« « » « × + »
  FIT [ENTER]
« « LN » « × EXP SWAP
  EXP × » FIT [ENTER]
« « SWAP LN SWAP »
  « LN × + » FIT [ENTER]
« « LN SWAP LN SWAP »
  « SWAP ^ SWAP EXP × »
  FIT [ENTER]
```



```
1: «  « LN SWAP LN SWAP
   »  « SWAP ^ SWAP EXP
   * » FIT »
```

Store the programs above. Note that the *LIN* program defines a null transform to the data. Your [USER] menu may differ from the one shown below.

```
[CLEAR] [USER]
'PWR [STO]
'LOGF [STO]
'EXPF [STO]
'LIN [STO]
```

```
3:
2:
1:
```

Key in the general curve fitting program below. The program uses the transforms defined above in the calling programs and returns the regression equation and coefficient of determination for a measure of the goodness of fit.

| **Program:** | **Comments:** |
|---|---|
| `« → xf1 xf2` | Store data transform and equation generator. |
| `« RCLΣ 'TMP' STO NΣ → n` | Save original data; get data count for looping. |
| `« 1 n START Σ- ARRY→ DROP xf1 EVAL NEXT` | Transform original data onto stack. |
| `n 2 2 →LIST →ARRY STOΣ »` | Put transformed data into current statistical matrix. |
| `LR DUP2 RCLΣ TRN RCLΣ ×` | Compute $A$ and $b$; duplicate for regression equation and $R^2$. |
| `{2 1} GET × TOT {2} GET SQ NΣ ÷ - SWAP TOT {2} GET × +` | Compute numerator of $R^2$. |
| `NΣ ÷ MEAN Σ+ VAR {2} GET ÷` | Compute denominator of $R^2$ ($n \cdot VAR_p(Y_i)$). |
| `TMP 'TMP' PURGE STOΣ` | Restore original data; purge temporary variable. |
| `ROT ROT 'X' xf2 EVAL STEQ RCEQ » »` | Generate regression equation and store. |

```
[ENTER] 'FIT [STO]
```

Key in the plotting program below. The program scales the plotting region by the statistical data and overlays the data and curve plots.

```
[PLOT]
« CLLCD SCLΣ DRWΣ DRAW
[ENTER]
```

```
2:
1:  « CLLCD SCLΣ DRWΣ
    DRAW »
```

Store the program in the variable *PLOT*.

```
'PLOT [STO]
```

```
3:
2:
1:
```

You may choose to enhance the program above by modifying the axes position according to the data set (for example, changing the axes position to the midpoints between the minimum and maximum points).

**Example:** Fit the following set of data into a straight line.

| $x_i$ | $y_i$ |
|---|---|
| ---- | ---- |
| 40.5 | 104.5 |
| 38.6 | 102 |
| 37.9 | 100 |
| 36.2 | 97.5 |
| 35.1 | 95.5 |
| 34.6 | 94 |

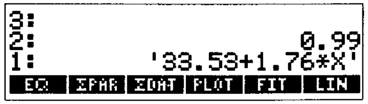
Clear the current statistical matrix and enter the data.

```
STAT  CLΣ
[40.5,104.5  Σ+
[38.6,102  Σ+
[37.9,100  Σ+
[36.2,97.5  Σ+
[35.1,95.5  Σ+
[34.6,94  Σ+
```


```
3:
2:
1:
```

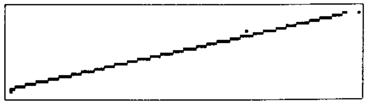
Compute the regression equation and coefficient of determination.

```
USER  LIN
```


```
3:
2:                0.99
1:     '33.53+1.76*X'
```

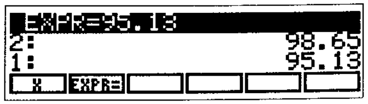
Plot the equation.

```
PLOT
```



Find estimates for $\hat{y}$ at $x = 37$ and $x = 35$.

```
ATTN  SOLV  SOLVR
37  X  EXPR=
35  X  EXPR=
```


```
EXPR=95.13
2:               98.65
1:               95.13
```

**Example:** Fit the following set of data into an exponential curve.

| $x_i$ | $y_i$ |
|---|---|
| ---- | ---- |
| .72 | 2.16 |
| 1.31 | 1.61 |
| 1.95 | 1.16 |
| 2.58 | .85 |
| 3.14 | 0.5 |

Clear the current statistical matrix and enter the data.

```
CLEAR
STAT  CLΣ
[.72,2.16  Σ+
[1.31,1.61  Σ+
[1.95,1.16  Σ+
[2.58,.85  Σ+
[3.14,.5  Σ+
```


```
3:
2:
1:
```

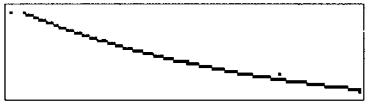
Compute the regression equation and coefficient of determination.

```
USER  EXPF
```


```
2:                 0.98
1:  'EXP(-(0.58*X))*3.45
    '
```

Plot the equation.

```
PLOT
```



Find estimates for $\hat{y}$ at $x = 1.5$ and $x = 2$.

```
ATTN  SOLV  SOLVR
1.5  X  EXPR=
2  X  EXPR=
```

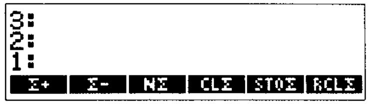

```
EXPR=1.08
2:                1.44
1:                1.08
```

**Example:** Fit the following set of data into a logarithmic curve.

| $x_i$ | $y_i$ |
|---|---|
| ---- | ---- |
| 3 | 1.5 |
| 4 | 9.3 |
| 6 | 23.4 |
| 10 | 45.8 |
| 12 | 60.1 |

Clear the current statistical matrix and enter the data.

```
CLEAR
STAT  CLΣ
[3,1.5  Σ+
[4,9.3  Σ+
[6,23.4  Σ+
[10,45.8  Σ+
[12,60.1  Σ+
```



```
3:
2:
1:
 Σ+ | Σ- | NΣ | CLΣ | STOΣ | RCLΣ
```

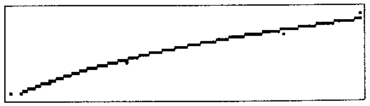
Compute the regression equation and coefficient of determination.

```
USER  LOGF
```

```
3:
2:                     0.98
1: '-47.02+41.39*LN(X)'
 FIT | LIN | EXPF | LOGF | PWR | GET1
```

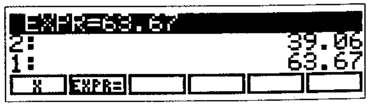
Plot the equation.

```
PLOT
```



Find estimates for $\hat{y}$ at $x = 8$ and $x = 14.5$.

```
ATTN  SOLV  SOLVR
8  X  EXPR=
14.5  X  EXPR=
```



```
EXPR=63.67
2:                    39.06
1:                    63.67
 X | EXPR= |  |  |  |
```

**Example:** Fit the following set of data into a power curve.

| $x_i$ | $y_i$ |
|---|---|
| ---- | ---- |
| 10 | .95 |
| 12 | 1.05 |
| 15 | 1.25 |
| 17 | 1.41 |
| 20 | 1.73 |
| 22 | 2.00 |
| 25 | 2.53 |
| 27 | 2.98 |
| 30 | 3.85 |
| 32 | 4.59 |
| 35 | 6.02 |

Clear the current statistical matrix and enter the data.

```
CLEAR
STAT  CLΣ
[10,.95  Σ+
[12,1.05  Σ+
[15,1.25  Σ+
[17,1.41  Σ+
[20,1.73  Σ+
[22,2  Σ+
[25,2.53  Σ+
[27,2.98  Σ+
[30,3.85  Σ+
[32,4.59  Σ+
[35,6.02  Σ+
```

```
3:
2:
1:
 Σ+ | Σ- | NΣ | CLΣ | STOΣ | RCLΣ
```

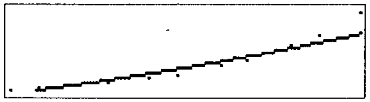
Compute the regression equation and coefficient of determination.

```
USER  PWR
```

```
3:
2:                     0.94
1:            'X^1.46*0.03'
 EXPF | LOGF | PWR | GET1 |  |
```

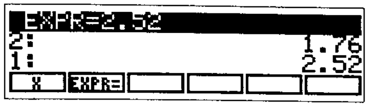
Plot the equation.

```
PLOT
```



Find estimates for $\hat{y}$ at $x = 18$ and $x = 23$.

```
ATTN  SOLV  SOLVR
18    X     EXPR=
23    X     EXPR=
```



```
EXPR=2.52
2:                   1.76
1:                   2.52
 X  EXPR=
```

## Multiple Regressions on the Same Data

Because the original, untransformed data is restored to the current statistical matrix, repeated and different regressions can be tried on the same data set. The equation plots can also be overlayed with a simple program like the one that follows.

**Example: Plotting Multiple Regressions.** For the data entered for the power curve fit in the preceding example, plot the curves for both power and exponential regressions, and compare their relative coefficients of determination.

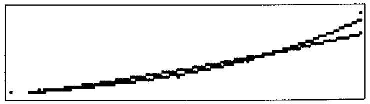
The program below performs the power curve fit, plots the data and curve, performs the exponential fit, and draws it.

```
CLEAR
« PWR PLOT EXPF DRAW
ENTER
```

```
2:
1: « PWR PLOT EXPF DRAW
   »
STEQ RCEQ PMIN PMAX INDEP DRAW
```

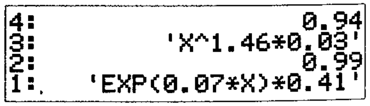
Execute the program. Note: You may find it necessary to purge unused variables to provide sufficient space in the HP-28C for both the curve fitting program and graphics display memory.

```
EVAL
```



Now compare the equations and the coefficients of determination.

```
ATTN  <>
```



```
4:                   0.94
3:         'X^1.46*0.03'
2:                   0.99
1:   'EXP(0.07*X)*0.41'
```

The exponential curve is a better fit.

Program *PLOT* is used in the Polynomial Regression section. If you wish, you can purge the other variables and programs created in the section.

---

## Multiple Linear Regression

This problem section provides a program for computing regression coefficients to a linear equation in two or three independent variables by the least squares method. The coefficient of determination is also computed, and point estimates based on the regression line can be computed.

### Two Independent Variables

For a set of data points $(x_i, y_i, t_i)$, the linear equation has the form

$$t = a + bx + cy$$

Regression coefficients $a$, $b$, and $c$ are calculated by solving the following system of equations.

$$\begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i \\ \sum y_i & \sum y_i x_i & \sum y_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum t_i \\ \sum x_i t_i \\ \sum y_i t_i \end{bmatrix}$$

The coefficient of determination is defined below.

$$R^2 = \frac{a\sum t_i + b\sum x_i t_i + c\sum y_i t_i - \frac{1}{n}(\sum t_i)^2}{\sum t_i^2 - \frac{1}{n}(\sum t_i)^2}$$

### Three Independent Variables

For a set of data points $(x_i, y_i, z_i, t_i)$, the linear equation has the form

$$t = a + bx + cy + dz$$

Regression coefficients $a$, $b$, $c$, and $d$ are calculated by solving the following system of equations.

$$\begin{bmatrix} n & \sum x_i & \sum y_i & \sum z_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i & \sum x_i z_i \\ \sum y_i & \sum y_i x_i & \sum y_i^2 & \sum y_i z_i \\ \sum z_i & \sum z_i x_i & \sum z_i y_i & \sum z_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} \sum t_i \\ \sum x_i t_i \\ \sum y_i t_i \\ \sum z_i t_i \end{bmatrix}$$

The coefficient of determination is defined below.

$$R^2 = \frac{a\sum t_i + b\sum x_i t_i + c\sum y_i t_i + d\sum z_i t_i - \frac{1}{n}(\sum t_i)^2}{\sum t_i^2 - \frac{1}{n}(\sum t_i)^2}$$

The following minimum condition for the number of data points $n$ must be satisfied:

- $n \geq 3$ for the case of two independent variables.
- $n \geq 4$ for the case of three independent variables.

---

## Multiple Linear Regression Program

The program below finds the regression line for both two and three independent variables. It also calculates $R^2$.

| **Program:** | **Comments:** |
|---|---|
| `« RCLΣ SIZE LIST→ DROP → m` | Begin to build the left-most matrix in the system of equations. $m$ is the number of elements in each data vector. |
| `« 1 →LIST 1 CON ARRY→ DROP` | Generate NΣ 1's on the stack. |
| `RCLΣ TRN ARRY→ LIST→ DROP SWAP 1 + SWAP 2 →LIST →ARRY` | Combine the 1's into the array. |
| `DUP TRN ×` | Generate the first row and column and covariance data. |
| `m m 1 + 2 →LIST RDM` | Drop the last row by redimensioning. |
| `TRN ARRY→ DROP m 1 2 →LIST →ARRY → rhs` | Pull out the right-hand-side of the system solution and save. |
| `« m m 2 →LIST →ARRY INV rhs ×` | Form the left-hand-side matrix, invert, and compute the regression coefficients. |
| `DUP TRN rhs × ARRY→ DROP` | Compute the first $m$ terms of the numerator of $R^2$. |
| `TOT m 1 →LIST GET SQ NΣ ÷ -` | Complete the numerator. |
| `NΣ ÷ MEAN Σ+ VAR m 1 →LIST GET ÷` | The denominator of $R^2$ is $n \cdot VAR_{population}$. |
| `Σ- DROP » » »` | Undo the change made to compute the population variance. |

`ENTER 'MLR STO`

**Example:** Find the regression coefficients and coefficient of determination for the following set of data.

| $x_i$ | $y_i$ | $z_i$ | $t_i$ |
|---|---|---|---|
| 7 | 25 | 6 | 60 |
| 1 | 29 | 15 | 52 |
| 11 | 56 | 8 | 20 |
| 11 | 31 | 8 | 47 |
| 7 | 52 | 6 | 33 |

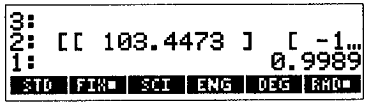
Clear the current statistics matrix and enter the data.

```
CLEAR
STAT  CLΣ
[7,25,6,60  Σ+
[1,29,15,52  Σ+
[11,56,8,20  Σ+
[11,31,8,47  Σ+
[7,52,6,33  Σ+
```

```
3:
2:
1:
 Σ+ | Σ- | NΣ | CLΣ | STOΣ | RCLΣ
```

Compute the regression coefficients and coefficient of determination.

```
USER  MLR
MODE  4  FIX
```

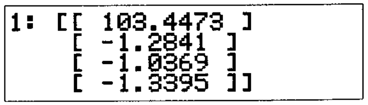


```
3:
2: [[ 103.4473 ]  [ -1…
1:                0.9989
 STD | FIX■ | SCI | ENG | DEG | RAD■
```

The coefficient of determination is 0.9989 .

Drop it and display the values for $a$, $b$, $c$, and $d$.

```
DROP  <>
```



```
1: [[ 103.4473 ]
    [ -1.2841 ]
    [ -1.0369 ]
    [ -1.3395 ]]
```

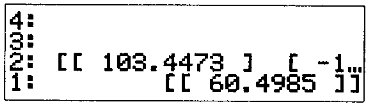
The regression line is $t = 103.4473 - 1.2841x - 1.0369y - 1.3395z$.

You can also compute estimates for $\hat{t}$ by multiplying the regression coefficients matrix by a matrix of values for the independent variables.

**Example:** Find $t$ for $x = 7$, $y = 25$, and $z = 6$, and $x = 1$, $y = 29$, and $z = 15$ for the problem above.

First make a copy of the coefficient matrix for the two computations of $\hat{t}$. Enter the first set of values for the independent variables. Note that a one is entered for the multiplication with the coefficient $a$.

```
ENTER
[[1,7,25,6  SWAP  ×
```

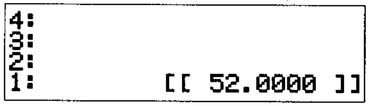


```
4:
3:
2: [[ 103.4473 ]  [ -1…
1:          [[ 60.4985 ]]
```

The estimate $\hat{t}$ is 60.4985 .

Compute $\hat{t}$ for the second set of values.

```
DROP
[[1,1,29,15  SWAP  ×
```



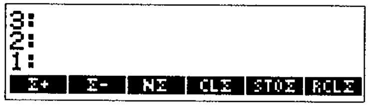
```
4:
3:
2:
1:          [[ 52.0000 ]]
```

**Example:** Find the regression line and the coefficient of determination for the following data.

| $x_i$ | $y_i$ | $t_i$ |
|---|---|---|
| 1.5 | 0.7 | 2.1 |
| 0.45 | 2.3 | 4.0 |
| 1.8 | 1.6 | 4.1 |
| 2.8 | 4.5 | 9.4 |

Clear the stack and the current statistics matrix; enter the data.

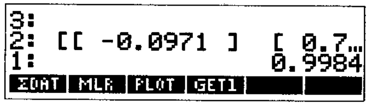
```
CLEAR  STAT  CLΣ
[1.5,.7,2.1  Σ+
[.45,2.3,4  Σ+
[1.8,1.6,4.1  Σ+
[2.8,4.5,9.4  Σ+
```



```
3:
2:
1:
 Σ+ | Σ- | NΣ | CLΣ | STOΣ | RCLΣ
```

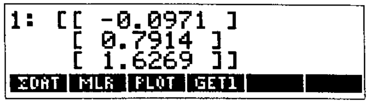
Find the regression line and coefficient of determination.

```
USER  MLR
```



```
3:
2: [[ -0.0971 ]  [ 0.7…
1:               0.9984
 ΣDAT | MLR | PLOT | GET1 |  |
```

DROP



```
1: [[ -0.0971 ]
    [ 0.7914 ]
    [ 1.6269 ]]
```

The regression line is $t = -.0971 + .7914x + 1.6269y$ . The same techniques described in the previous example may be used for computing $\hat{t}$.

Save programs *MLR* and *PLOT* for the Polynomial Regression section. Purge the other variables created in this section.

# Polynomial Regression

This problem section provides general programs for calculating the regression coefficients of parabolic and cubic equations for sets of paired data points using the least squares method. The coefficient of determination is also computed, and point estimates based on the regression equation can be computed.

## Parabolic Regression

For a set of data points $(x_i, y_i)$, the parabolic equation has the form

$$y = a + bx + cx^2$$

Regression coefficients $a$, $b$, and $c$ are calculated by solving the following system of equations.

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

The coefficient of determination is defined below.

$$R^2 = \frac{a\sum y_i + b\sum x_i y_i + c\sum x_i^2 y_i - \frac{1}{n}(\sum y_i)^2}{\sum y_i^2 - \frac{1}{n}(\sum y_i)^2}$$

## Cubic Regression

For a set of data points $(x_i, y_i)$, the cubic equation has the form

$$y = a + bx + cx^2 + dx^3$$

Regression coefficients $a$, $b$, $c$, and $d$ are calculated by solving the following system of equations.

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \sum x_i^3 y_i \end{bmatrix}$$

The coefficient of determination is defined below.

$$R^2 = \frac{a\sum y_i + b\sum x_i y_i + c\sum x_i^2 y_i + d\sum x_i^3 y_i - \frac{1}{n}(\sum y_i)^2}{\sum(y_i^2) - \frac{1}{n}(\sum y_i)^2}$$

The following minimum condition for the number of data points $n$ must be satisfied:

- $n \geq 3$ for Parabolic Regression.
- $n \geq 4$ for Cubic Regression.

## Polynomial Regression Programs

The programs below transform the data to a form that can be used in the multiple linear regression program on page 49. By modifying the statistical data inputs to the form $[x_i\ x_i^2\ y_i]$, the multiple linear regression program, *MLR*, computes the regression coefficients and coefficient of determination for parabolic regression. Similarly, by including an $x_i^3$ term, the *MLR* program computes the coefficients for a cubic regression.

**Program:**

```
« SWAP DUP SQ 3 ROLL
{3} →ARRY Σ+ »
```

**Comments:**

Form $[x_i\ x_i^2\ y_i]$ and accumulate in the current statistical matrix.

ENTER 'PARA STO

**Program:**

```
« SWAP DUP DUP SQ
SWAP 3 ^ 4 ROLL {4}
→ARRY Σ+ »
```

**Comments:**

Form $[x_i\ x_i^2\ x_i^3\ y_i]$ and accumulate in the current statistical matrix.

ENTER 'CUB STO

**Example:** Find the regression coefficients and coefficient of determination for the following set of data.

| $x_i$ | $y_i$ |
|---|---|
| --- | --- |
| .8 | 24 |
| 1 | 20 |
| 1.2 | 10 |
| 1.4 | 13 |
| 1.6 | 12 |

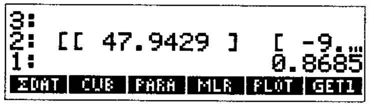
Clear the current statistical matrix and enter the data.

CLEAR
STAT CLΣ
USER
.8,24 CUB
1,20 CUB
1.2,10 CUB
1.4,13 CUB
1.6,12 CUB

```
3:
2:
1:
```

Compute the cubic regression coefficients and coefficient of determination.

MLR

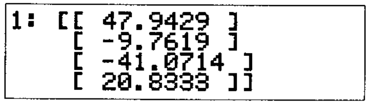


```
3:
2: [[ 47.9429 ]  [ -9.…
1:                0.8685
```

The coefficient of determination is 0.8685.

Drop it and display the values for $a$, $b$, $c$, and $d$.

DROP <>



```
1: [[ 47.9429 ]
    [ -9.7619 ]
    [ -41.0714 ]
    [ 20.8333 ]]
```

The regression equation is

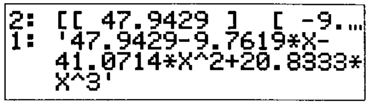
$$y = 47.9429 - 9.7619x - 41.0714x^2 + 20.8333x^3$$

You can also compute estimates for $\hat{y}$ by multiplying the regression coefficients matrix by a matrix of values for the independent variables.

**Example:** Find $\hat{y}$ for $x = 1$ and $x = 1.4$ for the problem above.

Enter the regression equation and use Solver to compute $\hat{y}$.

```
'47.9429-9.7619×X-
41.0714×X^2+20.8333×X^3
ENTER
```



```
2: [[ 47.9429 ]  [ -9.…
1: '47.9429-9.7619*X-
   41.0714*X^2+20.8333*
   X^3'
```

Store the equation and compute $\hat{y}$ for $x = 1$.

```
SOLV  STEQ  SOLVR
1  X
EXPR=
```

```
EXPR=17.9429
2: [[ 47.9429 ]  [ -9.…
1:              17.9429
```

Repeat for $x = 1.4$.

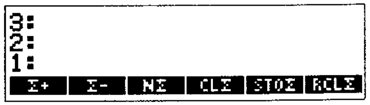
```
1.4  X
EXPR=
```

```
EXPR=10.9429
2:              17.9429
1:              10.9429
```

With the regression equation entered above, compare the original statistical data to a plot of the equation.

First, clear the current statistical matrix and enter the original data. (For larger matrices, the *GET1* and the ARRAY menu functions can be used to extract the first and last columns of data and construct the statistical matrix without reentering of the data.)

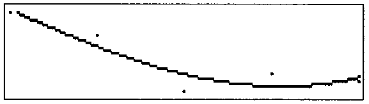
```
CLEAR
STAT  CLΣ
[.8,24  Σ+
[1,20  Σ+
[1.2,10  Σ+
[1.4,13  Σ+
[1.6,12  Σ+
```



```
3:
2:
1:
```

Use the *PLOT* program from the section entitled "Curve Fitting" to plot the data and the cubic equation.

```
USER  PLOT
```



**Example:** Find the parabolic equation and the coefficient of determination for the following data.

| $x_i$ | $y_i$ |
|---|---|
| --- | --- |
| 1 | 5 |
| 2 | 12 |
| 3 | 34 |
| 4 | 50 |
| 5 | 75 |
| 6 | 84 |
| 7 | 128 |

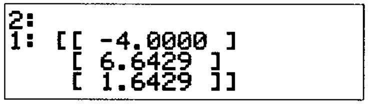
Clear the stack and the current statistical matrix; enter the data.

```
CLEAR  STAT  CLΣ
USER
1,5  PARA
2,12  PARA
3,34  PARA
4,50  PARA
5,75  PARA
6,84  PARA
7,128  PARA
```

```
3:
2:
1:
```

Find the parabolic regression equation and coefficient of determination.

```
MLR
DROP  <>
```



```
2:
1: [[ -4.0000 ]
    [ 6.6429 ]
    [ 1.6429 ]]
```

The regression line is $y = -4.0000 + 6.6429x + 1.6429x^2$. The same techniques described in the previous example may be used for computing $\hat{y}$.

# Test Statistics and Confidence Intervals

Decisions based on sample data can be directed with the use of test statistics. A variety of test statistics for different hypotheses and assumptions can be calculated with the HP-28S and HP-28C. This section presents three such test statistics – paired *t* statistic, *t* statistic for two means, and chi-square statistic. Additional test statistics for different hypotheses are readily computed with similar, simple procedures. The test statistics are used in conjunction with the upper-tail probability commands of the calculator to determine confidence intervals.

---

## Paired t Statistic

Given a set of paired observations from two normal populations with unknown means $\mu_1, \mu_2$

$$\begin{bmatrix} x_i & y_i \\ \text{---} & \text{---} \\ x_1 & y_1 \\ x_2 & y_2 \\ \cdots & \cdots \\ x_n & y_n \end{bmatrix}$$

the test statistic

$$t = \frac{\overline{D}}{s_D} \cdot \sqrt{n}$$

with $n-1$ degrees of freedom can be used to test the null hypothesis

$$H_0: \mu_1 = \mu_2$$

The variable definitions are

$$D_i = x_i - y_i$$

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$$

and

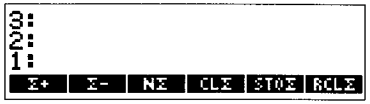
$$s_D = \left( \frac{\sum D_i^{\,2} - \frac{1}{n}(\sum D_i)^2}{n-1} \right)^{\frac{1}{2}}$$

**Example:** Test the null hypothesis that $\mu_1 = \mu_2$ for the following data pairs.

| $x_i$ | $y_i$ |
|---|---|
| --- | --- |
| 14 | 17 |
| 17.5 | 20.7 |
| 17 | 21.6 |
| 17.5 | 20.9 |
| 15.4 | 17.2 |

Clear the current statistical matrix and key in the data.

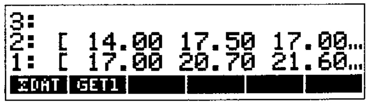
```
CLEAR  MODE  2  FIX

STAT  CLΣ
[14,17  Σ+
[17.5,20.7  Σ+
[17,21.6  Σ+
[17.5,20.9  Σ+
[15.4,17.2  Σ+
```



```
3:
2:
1:
```

Use *GET1* from the USER menu to recover the two columns of data. Remember, your USER menu may not appear identical to the menu shown below.

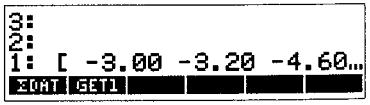
```
1  USER  GET1
2  GET1
```



```
3:
2: [ 14.00 17.50 17.00…
1: [ 17.00 20.70 21.60…
```

Compute the difference of the data pairs.

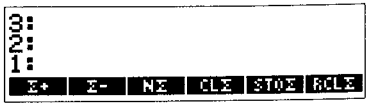
```
-
```



```
3:
2:
1: [ -3.00 -3.20 -4.60…
```

Redimension and store the difference matrix as the current statistical matrix. (Save the original data if desired.)

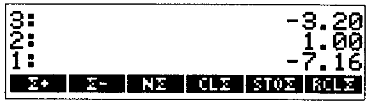
```
STAT  NΣ  1  2  LIST  →LIST
ARRAY  RDM  STAT  STOΣ
```



```
3:
2:
1:
```

Compute the *t* statistic.

```
MEAN  SDEV
STACK  DUP2
÷  STAT  NΣ  √  ×
```



```
3:                    -3.20
2:                     1.00
1:                    -7.16
```

The mean $\overline{D}$ is −3.20 .
$s_D$ is 1.00 .
*t* is −7.16 .
The degrees of freedom are 4.00 .

**Example:** Determine if the hypothesis $H_0$ of the previous problem should be rejected at a 0.05 level of significance.

First create a program for a general solution to the Student's *t* distribution as described in the STAT section of the reference manual.

```
CLEAR
« P N X UTPT - »
ENTER
```

```
3:
2:
1:       « P N X UTPT - »
```

Store it for use in the Solver.

```
SOLV  STEQ  SOLVR
```

```
3:
2:
1:
```

Enter the degrees of freedom and the level of significance. Note that for a two-tailed test at a 0.05 level of significance, you compute the value for the confidence interval $-t_{.975}$ to $t_{.975}$. (Remember, the keys ☐ X indicate the shift key followed by the user-defined SOLVR key X. Pressing these keys tells the Solver to seek a solution for "X" in the specified equation.)

```
4  N  .025  P
☐  X
```

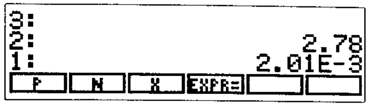
```
X: 2.78
Sign Reversal
1:                     2.78
```

Thus the hypothesis is rejected since *t* falls outside the range (−2.78,2.78).

**Example:** Compute the level of significance for which the hypothesis $H_0$ will be accepted.

Enter the $t$ statistic computed in the first example. Note that the absolute value is input, corresponding to the upper-tail portion of the probability function.

7.16 X
☐ P
2 ×



```
3:
2:                2.78
1:            2.01E-3
 P   N   X  EXPR=
```

The probability is multiplied by 2 for the upper- and lower-tails of the probability function outside the range (−7.16,7.16).

Rather than using the program from the previous example, you can also compute the level of significance directly with the UTPT command. The keystrokes

4,7.16 STAT UTPT 2 ×

generate the same result as above.

Exit from the Solver menu and purge the variables created in this section.

SOLV {'P''N''X''EQ PURGE

## t Statistic for Two Means

Suppose $\{x_1, x_2, \cdots, x_{n1}\}$ and $\{y_1, y_2, \cdots, y_{n2}\}$ are two independent random samples from two normal populations with unknown means $\mu_1$ and $\mu_2$ and the same unknown variance $\sigma^2$.

The null hypothesis

$$H_0 \colon \mu_1 - \mu_2 = d$$

can be tested with the $t$ statistic

$$t = \frac{\overline{x} - \overline{y} - d}{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{\frac{1}{2}} \left(\frac{\sum x_i^2 - n_1\overline{x}^2 + \sum y_1^2 - n_2\overline{y}^2}{n_1 + n_2 - 2}\right)^{\frac{1}{2}}}$$

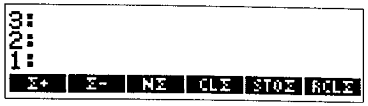
This $t$ statistic has the $t$ distribution with $n_1 + n_2 - 2$ degrees of freedom for testing the null hypothesis $H_0$.

**Example:** Test the null hypothesis that $H_0 \colon \mu_1 = \mu_2$ (i.e. $d=0$) for the data below.

| $x_i$ | $y_i$ |
|---|---|
| --- | --- |
| 79 | 91 |
| 84 | 103 |
| 108 | 90 |
| 114 | 113 |
| 120 | 108 |
| 103 | 87 |
| 122 | 100 |
| 120 | 80 |
|  | 99 |
|  | 54 |

Clear the current statistical matrix and accumulate the $x$ data.

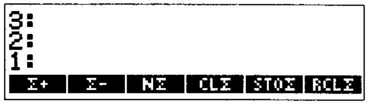
```
CLEAR
STAT  CLΣ
79   Σ+
84   Σ+
108  Σ+
114  Σ+
120  Σ+
103  Σ+
122  Σ+
120  Σ+
```



```
3:
2:
1:
```

Compute the mean, variance, and number of data points, then store these values.

```
MEAN  'MX  STO
VAR   'VX  STO
NΣ    'NX  STO
```



```
3:
2:
1:
```

Clear the current statistical matrix and accumulate the $y$ data.

```
CLΣ
91   Σ+
103  Σ+
90   Σ+
113  Σ+
108  Σ+
87   Σ+
100  Σ+
80   Σ+
99   Σ+
54   Σ+
```

```
3:
2:
1:
```

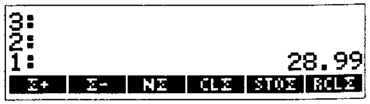
Compute the $t$ statistic. First compute the numerator. Recall that $d=0$.

```
USER  MX  STAT  MEAN  -
```

```
3:
2:
1:                  13.75
```

Compute the first part of the denominator and divide.

```
USER  NX  1/x
STAT  NΣ  1/x
+  √  ÷
```



```
3:
2:
1:                  28.99
```

Compute the second part of the denominator and divide.

```
NΣ  1  -
VAR  ×
USER  NX  1  -
VX  ×  +
NX  STAT  NΣ  +  2  -
÷  √  ÷
```

```
3:
2:
1:                   1.73
```

The $t$ statistic is 1.73 with 16 degrees of freedom.

**Example:** Compute the level of significance for the two-tailed test on the range (−1.73,1.73) from the preceding example.

The display shown here includes COMB and PERM, two functions that do not appear on the HP-28C STAT display. The appendix to this book describes HP-28C programs that provide these functions.

```
16,1.73  UTPT
2  ×
```

```
3:
2:                   1.73
1:                   0.10
```

The hypothesis cannot be rejected at, or below, this level of significance.

You may also choose to test the assumption made in this problem section that the unknown variances are equal. For this purpose, compute the F statistic

$$\frac{s_{\max}^2}{s_{\min}^2}$$

$\frac{s_1^2}{s_2^2}$ has $n_1-1, n_2-1$ degrees of freedom.

$\frac{s_2^2}{s_1^2}$ has $n_2-1, n_1-1$ degrees of freedom.

$s_{\max}^2$ is the maximum of the sample variances $s_1^2$ and $s_2^2$.

$s_{\min}^2$ is the minimum of the two sample variances.

You can then compute the level of significance with the UTPF command, using the same approach as the example above.

Purge the variables created in this section.

```
{'NX''VX''MX [PURGE]
```

## Chi-Square Statistic

This section provides a simple program for computing the $\chi^2$ statistic for the goodness of fit test.

The equation computed is

$$\chi_1{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

with $n - 1$ degrees of freedom.

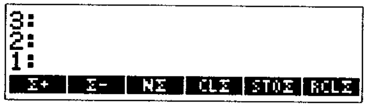
$O_i$ is the observed frequency.
$E_i$ is the expected frequency.
$n$ is the number of classes.

**Example:** Find the value of the chi-square statistic for the goodness of fit for the following data set.

| $O_i$ | $E_i$ |
|---|---|
| 8 | 9.6 |
| 50 | 46.75 |
| 47 | 51.85 |
| 56 | 54.4 |
| 5 | 8.25 |
| 14 | 9.15 |

Clear the current statistical matrix and enter the data.

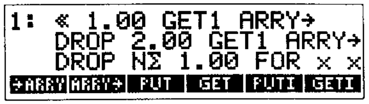
```
[CLEAR]
[STAT] [CLΣ]
[8,9.6 [Σ+]
[50,46.75 [Σ+]
[47,51.85 [Σ+]
[56,54.4 [Σ+]
[5,8.25 [Σ+]
[14,9.15 [Σ+]
```



```
3:
2:
1:
```

Enter the program below to compute the chi-square statistic. The program has the same form as the delta percent program from page 32. You may wish to refer to it for comments regarding the approach used for this program. Note that this program differs by specifying columns one and two of the statistical matrix to calculate the result. The program could be generalized for any pair of columns by retrieving the contents of Σ*PAR* for the column specification.

```
« 1 GET1 ARRY→ DROP 2
GET1 ARRY→ DROP NΣ 1
FOR x x NΣ + ROLL → ofq
« NΣ ROLL → efq
« '(ofq-efq)^2÷efq'
EVAL » » -1 STEP NΣ
1 2 →LIST →ARRY CNRM »
ENTER
```



```
1: « 1.00 GET1 ARRY→
   DROP 2.00 GET1 ARRY→
   DROP NΣ 1.00 FOR x x
```

Save the program in the variable *CHI* for repeated use.

`'CHI STO`

```
3:
2:
1:
```

Execute the program.

USER CHI

```
3:
2:
1:                 4.84
```

The chi-square statistic $\chi^2$ is 4.84 with 5 degrees of freedom.

**Example:** Compute the level of significance for the example above.

Enter the degrees of freedom and the $\chi^2$ statistic, and compute the upper-tail probability for the $\chi^2$ distribution. (The display shown here includes COMB and PERM, two functions that do not appear on the HP-28C STAT display. The appendix to this book describes HP-28C programs that provide these functions.)

`5,4.84` STAT UTPC

```
3:
2:                 4.84
1:                 0.44
```

The hypothesis cannot be rejected at, or below, this level of significance.

# Appendix: Combinations and Permutations

The HP-28C programs that follow provide simple building blocks for combinatorial analysis. Complex problems are readily evaluated by combining the results left on the stack or by using the programs below as subroutines.

HP-28S users, please note: the COMB and PERM programs defined here appear on the STAT menu of your calculator. If you wish to perform the examples that follow the programs, change your display to the STAT menu and use the PERM and COMB menu items as described in the examples.

## Permutations

Given *X* distinct objects, the number of ways to select and arrange *Y* of these objects in different order is computed by the formula below.

$$_XP_Y = \frac{X!}{(X-Y)!}$$

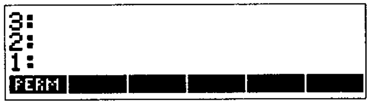
Clear the stack and key in the permutations program.

```
CLEAR MODE STD USER
« → x y 'FACT(x)÷
FACT(x-y ENTER
```

```
2:
1: « → x y 'FACT(x)/
   FACT(x-y)' »
```

Store the program in the variable *PERM*.

`'PERM STO`



```
3:
2:
1:
```

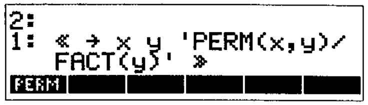
## Combinations

Combinations ignore the order in the $Y$ objects chosen and are computed by the formula below.

$${}_XC_Y = \frac{X!}{Y!(X-Y)!}$$

Key in the combinations program.

```
« → x y 'PERM(x,y)÷
FACT(y ENTER
```



```
2:
1: « → x y 'PERM(x,y)/
   FACT(y)' »
PERM
```

Store the program in the variable *COMB*.

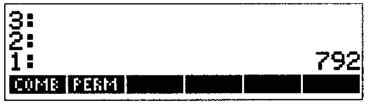
```
'COMB STO
```

```
3:
2:
1:
COMB PERM
```

**Example:** Compute how many five-person basketball squads can be formed from 12 players. The computation to be made is ${}_{12}C_5$.

With the program *COMB* keyed in as above, key in the parameters and evaluate the formula.

```
12 ENTER
5 COMB
```



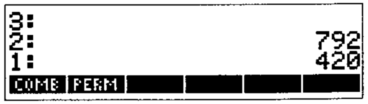
```
3:
2:
1:                 792
COMB PERM
```

792 squads can be formed. Any combination of five players is acceptable, since the combination program was used to compose the number of teams.

**Example:** For the problem above, what if one of the two tallest players *must* be on the squad, and these two players never play at the same time?

There are now ten players from which to select the four remaining positions, and two ways to select the fifth. Thus, compute ${}_{10}C_4 \cdot 2$.

```
10 ENTER
4 COMB
2 ×
```

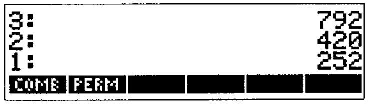


```
3:
2:                 792
1:                 420
COMB PERM
```

**Example:** Compute the number of options lost if both tall players from the previous example foul out.

Form the five-person squad from the remaining players.

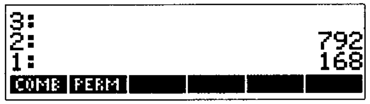
```
10 ENTER
5 COMB
```



```
3:                 792
2:                 420
1:                 252
COMB PERM
```

The options lost are computed by subtracting.

```
-
```

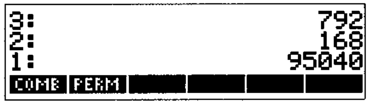


```
3:
2:                 792
1:                 168
COMB PERM
```

168 squad combinations were lost as a result.

**Example:** Compute the number of permutations of the twelve original players that are possible.

```
12 ENTER
5 PERM
```



```
3:                 792
2:                 168
1:               95040
COMB PERM
```

For large values of $X$ and $Y$, it may be desirable to use a program that computes the value of the combination or permutation formula by explicitly multiplying the appropriate terms of the factorials. This can improve the accuracy of the result.

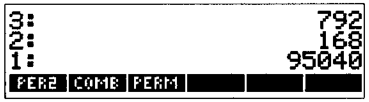
For example, rather than evaluating ${}_XP_Y$ by $FACT(X) \div (FACT(X-Y)$, compute it as the product of the appropriate terms: $X \cdot (X-1) \cdot (X-2) \cdots (X-Y-1)$.

Key in the following program and compute ${}_{12}P_5$ from the previous example.

```
« → x y « x y - 'y'
STO x WHILE x 1 - y >
REPEAT x 1 - DUP 'x'
STO × END ENTER
```

```
1: « → x y « x y - 'y'
   STO x WHILE x 1 - y
   > REPEAT x 1 - DUP
COMB PERM
```

```
'PER2 STO
```



```
3:                 792
2:                 168
1:               95040
PER2 COMB PERM
```

12 ENTER
5 PER2

```
3:                168
2:              95040
1:              95040
PER2 COMB PERM
```

An example of the accuracy difference between the two approaches can be seen by computing ${}_{20}P_{10}$.

If you wish, purge the programs and variables created in this section.

`{'COMB''PERM''PER2'` PURGE

## More Step-by-Step Solutions for Your HP-28S or HP-28C Calculator

These additional books offer a variety of examples and keystroke procedures to help set up your calculations the way *you* need them.

Practical routines show you how to use the built-in menus to solve problems more effectively, while easy-to-follow instructions help you create personalized menus.

### Algebra and College Math (00028-90101)

- Solve algebraic problems: polynomial long division, function evaluation, simultaneous linear equations, quadratic equations, logarithms, polynomial equations, matrix determinants, and infinite series.
- Perform trigonometric calculations: graphs of functions, relations and identities, inverse functions, equations, and complex numbers.
- Solve analytic geometry problems: rectangular and polar coordinates, straight line, circle, parabola, ellipse, hyperbola, and parmetric equations.

### Calculus (00028-90102)

- Perform function operations: definition, composition, analysis, angles between lines, and angles between a line and a function.
- Solve problems of differential calculus: function minimization, computing tangent lines and implicit differentiation.
- Obtain symbolic and numerical solutions for integral calculus problems: polynomial integration, area between curves, arc length of a function, surface area, and volume of a solid of revolution.

### Vectors and Matrices (00028-90105)

- Perform general matrix operations: summation, multiplication, determinant, inverse, transpose, conjugate, and minor rank.
- Solve a system of linear equations.
- Calculate several important vector operations.

- Learn methods for calculating eigenvalues and eigenvectors.
- Perform the method of least squares and Markov Chain calculations.

*And Specifically for Your HP-28S...*

**Electrical Engineering** (00028-90103)

- Use ready-to-go utilities for quick results.
- Calculate mesh currents and nodal voltages of DC or AC circuits.
- Display and optionally print a Bode diagram of a transfer function.
- Calculate the impedance of a passive one-port network.
- Convert impedance values between delta and Y networks.
- Perform ladder network analysis, truth table generation, finite Fourier transforms, and voltage/current division.

**Mathematical Applications** (00028-90111)

- Find the area and all sides and angles of any plane triangle.
- Perform synthetic division on polynomials of arbitrary order.
- Calculate all the roots of a first, second, third and fourth degree polynomial, with real or complex coefficients.
- Solve first and second-order differential equations.
- Convert the coordinates of two or three-dimensional vectors between two coordinate systems, where one system is translated and/or rotated with respect to the other.
- Collect statistical data points, and fit curves to the data.

## How to Order...

For the location and number of the U.S. dealer nearest you, call toll-free 1-800-752-0900. To order a book your dealer does not carry, call toll-free 1-800-538-8787 and refer to call code P270. Master Card, Visa, and American Express cards are welcome. For countries outside the U.S., contact your local Hewlett-Packard sales office.