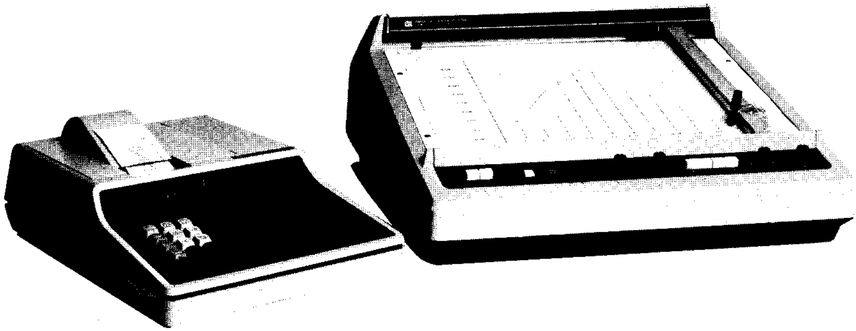
MODEL 5

HEWLETT-PACKARD
9805A STATISTICS CALCULATOR
EXPANDED STAT OPERATING GUIDE

Operating Guide



Expanded Stat System

JOHN KEITH

HEWLETT–PACKARD CALCULATOR PRODUCTS DIVISION
P.O. Box 301, Loveland, Colorado 80537, Tel. (303) 667-5000
Rue du Bois-du-Lan 7, CH-1217 Meyrin 2, Geneva, Tel. (022) 41 54 00

Table of Contents

Table of Contents

Chapter 1

Getting Started

Introduction

The Expanded Stat Programs Block adds these calculating and plotting functions to your 9805 Statistics Calculator System:

- **Program 1:** Automatically scales the plotter, then draws and labels axes – all to your specified dimensions.
- **Program 2:** Calculates power, exponential, and logarithmic regressions (curve fits), in addition to the linear and parabolic regressions already available. If your system has a plotter, you can also plot the 'best-fit model', since this program will plot any or all of the five regressions.
- **Program 3:** Calculates a one-way analysis of variance for any number of data groups.
- **Program 4:** Automatically plots a normal curve overlay using histogram data you've entered.

As the above list indicates, the Expanded Stat Block is designed to complement your system's plotting capability; however, programs 2 and 3 are easily run without the need of a plotter.

Running Programs

To run a program you must first press **START PRGM** (this readies the calculator for program operation); then after the paper advances, key in the program number and press **RUN/STOP.**

NOTE
Since the **AUTO·** feature cannot be used while running these programs, be sure the AUTO· key is up (switched off)!

To halt program operation, hold the **RUN/STOP** key down until the display returns (and the busy light goes out).

Chapter 2 offers complete instructions for running each program. Also, see the Statistics Calculator Pocket Reference for abbreviated instructions on running these programs.

> **CAUTION**
>
> DO NOT REMOVE THE PROGRAM BLOCK FROM THE CALCULATOR, since doing so may damage your calculator or program block!

Operating Notes

If a 'NOTE' is printed while you're running a program, refer to the list of operating notes at the back of this guide.

> **IMPORTANT!**
>
> If '2 NOTE 50' is printed after you've pressed a key do not attempt to do further plotting operations, since the note indicates that plotter control has been interrupted. To regain plotter control, you must switch the calculator OFF and then ON.

Service

If you suspect that a program is not operating properly, contact the nearest -hp- Sales and Service Office for assistance; office locations are listed at the back of the Statistics Calculator Operating Guide.

Chapter 2
Operating Instructions

Program 1: Plotting And Labeling Axes

This program is used to calibrate the plotter to user-specified dimensions.

General Instructions

1. Start Program 1.

2. Enter the following axes dimensions and tic mark intervals (press **RUN/STOP** after each entry):

   X_{min} = Minimum value for the X axis.

   X_{max} = Maximum value for the X (horizontal) axis.

   Y_{min} = Minimum value for the Y axis.

   Y_{max} = Maximum value for the Y (vertical) axis.

   X_{tic} = Tic interval for the X axis.

   Y_{tic} = Tic interval for the Y axis.

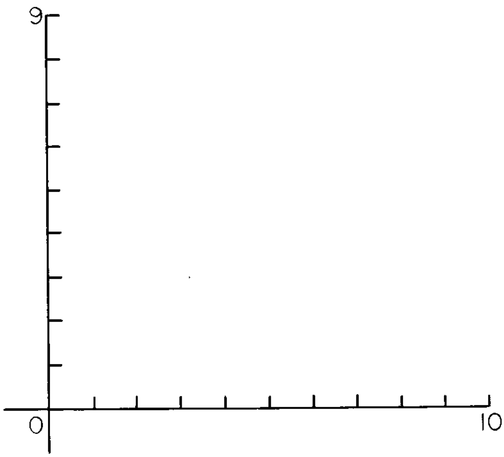
3. After the X and Y dimensions are entered, the axes are automatically plotted and the program stops to allow you to set the printing format before each axis is labeled (to set 'RND 1' format just press **RUN/STOP**).

To set this format:	Press:
'RND 0' through 'RND 6' 4.00 (RND 0) 0.625 (RND 3)	SHIFT PRINT 1 RND ( ) Digit LAST ENTRY RUN/STOP
Scientific Notation 1.0 E02 (SCI 1) 3.750 E00 (SCI 3)	SHIFT PRINT 1 RND ( ) • Digit LAST ENTRY RUN/STOP

If you do not wish to label the axes, press (START PRGM) to terminate the program.

Example:

Calibrate the plotter to the following dimensions; then plot and label each axis.



Here are the dimensions to enter for this example:

$X_{min} = 0$
$X_{max} = 10$
$Y_{min} = 0$
$Y_{max} = 9$
$X_{tic} = 1$
$Y_{tic} = 1$

Before starting the program, set up the plotter with paper and pen; then adjust the Graph Limits controls (set **Lower Left** controls first) for the desired size plot.

To start the program,

Press: [START PRGM]

CLEAR

Then Press: [1] [RUN STOP]

PRGM
1.00 ¤

To specify the X-axis and Y-axis dimensions, **press:**

Press: 0 [RUN STOP] 10 [RUN STOP] 0 [RUN STOP] 9 [RUN STOP]

```
  .00 ¤
10.00 ¤
  .00 ¤
 9.00 ¤
 1.00 ¤
 1.00 ¤
```

To specify X and Y tic mark intervals,

Press: 1 [RUN STOP] 1 [RUN STOP]

```
   AI =
  .00 ¤
  .00 ¤
```

After the axes are plotted, the program stops to allow you to set the printing format before the X axis is labeled. To label the X axis using 'ROUND 1' format,

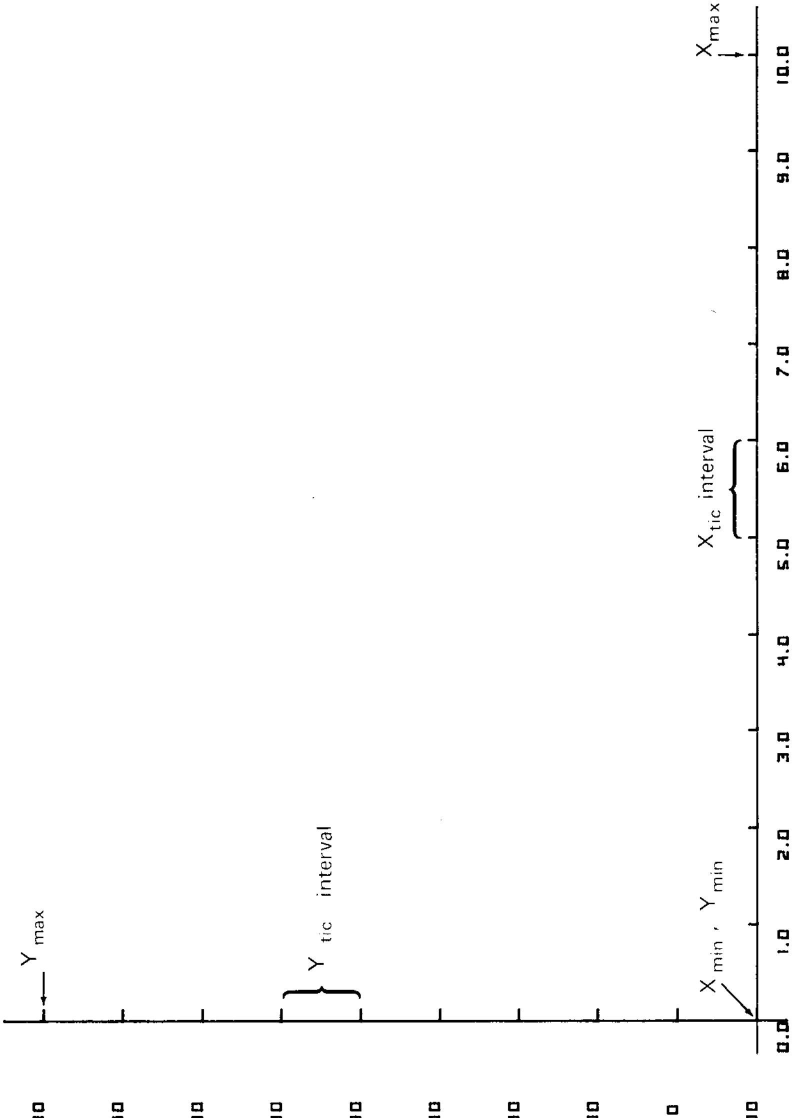
Press: [RUN STOP]

After the X axis is labeled, the program stops again. To label the Y axis using 'ROUND 2' format,

Press: SHIFT [PRINT] [1] [RND ( )] 2 [LAST ENTRY] [RUN STOP]

The final plot should look like the sample on page 6. Once the program's complete (indicated by advancing paper) the plotter is calibrated to your specified dimensions – you may use the labeled axes for plotting data points, histograms, regressions, etc.

For further examples using this program, see 'Program 2: Family Regression' and 'Program 4: Normal Curve Overlay'.



Program 1: Plotting and Scaling Axes

Labeling Histograms

Normally, the X-axis dimensions must be entered to specify the number of cells to be plotted and the width of each cell. For many applications, however, the histogram would be easier to read if the X axis were labeled to indicate *cell boundaries.*

Here's a suggested routine for using Program 1 to label a histogram with *cell boundaries* rather than *cell numbers*:

1. Determine the number of cells to be plotted; then determine the boundry of each cell. For example, if we were to generate a histogram of scores on an examination when the actual scores ranged from 5 to 98, there would be 10 cells and the cell boundries would be 10, 20, 30, . . . 100.

2. Using Program 1, plot and label the axes – specify an X_{min} value that corresponds to the lower boundary of the first cell and specify an X_{max} value that's slightly greater than the upper boundary of the last cell.

   For our example, enter these dimensions:

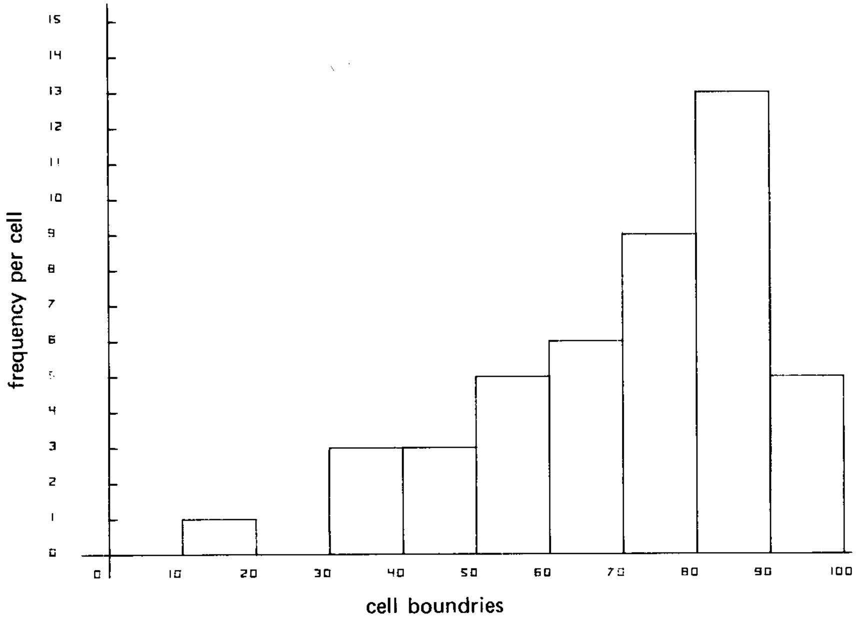
   $X_{min} = 0$
   $X_{max} = 100$
   $Y_{min} = 0$
   $Y_{max} =$ (highest frequency for any cell).
   $X_{tic} = 10$
   $Y_{tic} = 1$

3. After the axes are plotted and labeled (the X axis is now labeled according to cell boundaries), re-run Program 1, but this time specify X-axis dimensions that indicate the *number of cells* to be plotted. Then, when the program stops before labeling the axes, press **START PRGM** to terminate the program.

For our example, we would enter these dimensions:

$X_{min} = 0$
$X_{max} = 10$
Y_{min}
Y_{max} ← same as before
$X_{tic} = 0$
$Y_{tic} = 0$

4. Now plot the histogram in the usual manner. Here's a sample plot showing the X axis labeled by the method just described.



Program 2: Family Regression (Curve Fitting)

This program allows the user to fit five different regression models to a set of paired observations. If desired, data points can be plotted as they are input, and the selected lines can be plotted after the coefficients have been calculated. The r^2 value is also calculated for each model.

Regression analysis is a very useful tool when the object of an experiment is to predict the y (dependent variable) when a value of the x (independent variable) is given. The r^2 value (coefficient of determination) indicates the proportion of the total y-variation which is accounted for by the regression model.

The program allows you to calculate and plot any of these models:

Model 1: (Exponential) $y = ae^{bx}$
Model 2: (Power) $y = ax^b$
Model 3: (Logarithmic) $y = a + b \ln x$
Linear Model: $y = a + bx$
Parabolic or Quadratic Model: $y = a + bx + cx^2$

General Instructions

To perform a Family Regression:

> **NOTE**
>
> If you do not wish to use the plotter, do only the steps shown in green.

1. Set up the axes by using either Program 1 (plotting and labeling axes) or the **AXES** and **DATA ENTRY** keys. Make certain that both X_{max} and Y_{max} are *larger* than their largest respective values to be entered!

2. Start Program 2.

3. If you wish to plot the data points as they are entered, first select a plotting character by pressing the **CHAR#** key and keying in a character code (see the list of character codes in Appendix A), then press [AXES] [0]

4. To enter each pair of data points, first key in the x value (value of the independent variable for the data pair) and press the **RUN/STOP** key. Then key in the y value (value of the dependent variable for the data pair) and press the **RUN/STOP** key. Each set of data points is plotted (if so specified in Step 3) after it's entered. After each successive data pair is entered, the current number of entered data pairs is displayed. **IMPORTANT:** *Do not press DATA ENTY when entering points for this program.*

   To delete a data point, first press **STORE** and then enter the unwanted data values as described above. Also, if the data points are being plotted as they're entered, the plotter will 'X' out each deleted point after it's entered.

NOTE
Since calculations for Models 1, 2, and 3 include logarithms, *enter only values greater than 0.* If NOTE 9 or NOTE 12 occur, delete all values less than or equal to 0 and try again.

5. After all data points have been entered (and plotted),

   Press: [LAST ENTRY] [RUN/STOP]

   The program is now waiting for you to specify the model to be calculated (and plotted).

6. Before specifying the regression model, you may wish to change the plotting format (the solid-line format is automatically set before each model is calculated).

   To change the plotting format,

   Press: [SHIFT] [PRINT] [1] [)] [plot code] [LAST ENTRY]

Here's a list of plotting codes and the kind of plot specified:

Plotting Code	Description	Example Plot
0	solid line (automatically set)	————————
1	dashed line	— — — — — —
2	alternating long and short dashes	——— — ———
3	alternating dashes and dots	—— . . —— . . ——

7. To calculate the regression model:

Press: 1 RUN/STOP (exponential: $y = ae^{bx}$)

or Press: 2 RUN/STOP (power: $y = ax^b$)

or Press: 3 RUN/STOP (logarithmic: $y = a + b \ln x$)

or Press: LINEAR (linear: $y = a + bx$)

or Press: PARA (parabolic: $y = a + bx + cx^2$)

> **NOTE**
>
> If you wish to add or delete data points, first press **0,RUN/STOP** (instead of specifying a regression model) and then return to Step 4.

8. If more models are to be calculated, return to Step 7; otherwise,

Press: LAST ENTRY RUN/STOP

When the paper advances, the program is finished.

Example

Find the 'best-fit' model for the following five data points:

X	1	3	5	7	9
Y	10	6	4	3	1

1. Set up the plotter axes using Program 1.

Press: [START PRGM]

and then press [1] [RUN STOP]

Press: 0 [RUN STOP] 10 [RUN STOP] 0 [RUN STOP]

Press: 10 [RUN STOP] 1 [RUN STOP] 1 [RUN STOP]

```
          CLEAR

PRGM
  1.00         #

   .00         #
 10.00         #
   .00         #
 10.00         #
  1.00         #
  1.00         #
```

After the program stops,

Press: [RUN STOP]

After the program stops again, press [RUN STOP]

2. Once Program 1 is complete (axes are plotted and labeled), start the family regression program:

Press: [START PRGM]

Then press: [2] [RUN STOP]

```
          CLEAR

PRGM
  2.00         #

      DATACLEAR

V2    •••••••
```

3. To specify plotter character '1',

Press: (CHAR #) (1) — CT1

Then press: (AXES) (0)

4. To enter the data points (each point is plotted after its y value is entered),

Press		Printout
1 (RUN/STOP)	10 (RUN/STOP)	1.00 # 10.00 #
3 (RUN/STOP)	6 (RUN/STOP)	3.00 # 6.00 #
5 (RUN/STOP)	4 (RUN/STOP)	5.00 # 4.00 #
7 (RUN/STOP)	3 (RUN/STOP)	7.00 # 3.00 #
9 (RUN/STOP)	1 (RUN/STOP)	9.00 # 1.00 #

5. After all data points are entered,

Press: (LAST ENTRY) (RUN/STOP) — MODL =

NOTE

Before calculating each model, you may wish to change the plotting format by pressing the appropriate sequence of keys shown on Page 10.

6. To calculate the linear model,

Press: (LINEAR)

```
            A      =
      10.05        #
            B      =
-      1.05        #
r2                 =
        .94        #
```

To plot the linear model, **press** (PLOT)

7. To calculate the parabolic model:

Press: (PARA)

```
            A      =
      11.57        #
            B      =
-      1.94        #
            C      =
        .09        #
r2                 =
        .98        #
```

To plot the parabolic curve,

Press: (PLOT)

Notice that MODL = isn't printed after linear or parabolic calculations, eventhough the program is waiting to calculate another model.

8. To calculate and plot the exponential curve (model 1).

Press: (1) (SHIFT) ((CHS) (RUN STOP)

```
-      1.00        #
r2     A      B    =
        .94        #
      14.02        #
-       .26        #

        MODL       =
```

9. To calculate and plot the power curve (model 2),

Press: [2] [SHIFT] [(] (CHS) [RUN/STOP]

```
-        2.00      #
r²       A      B  =
          .79      #
        12.54      #
-         .89      #

        MODL       =
```

10. To calculate and plot the logarithmic model (model 3),

Press: [3] [SHIFT] [(] (CHS) [RUN/STOP]

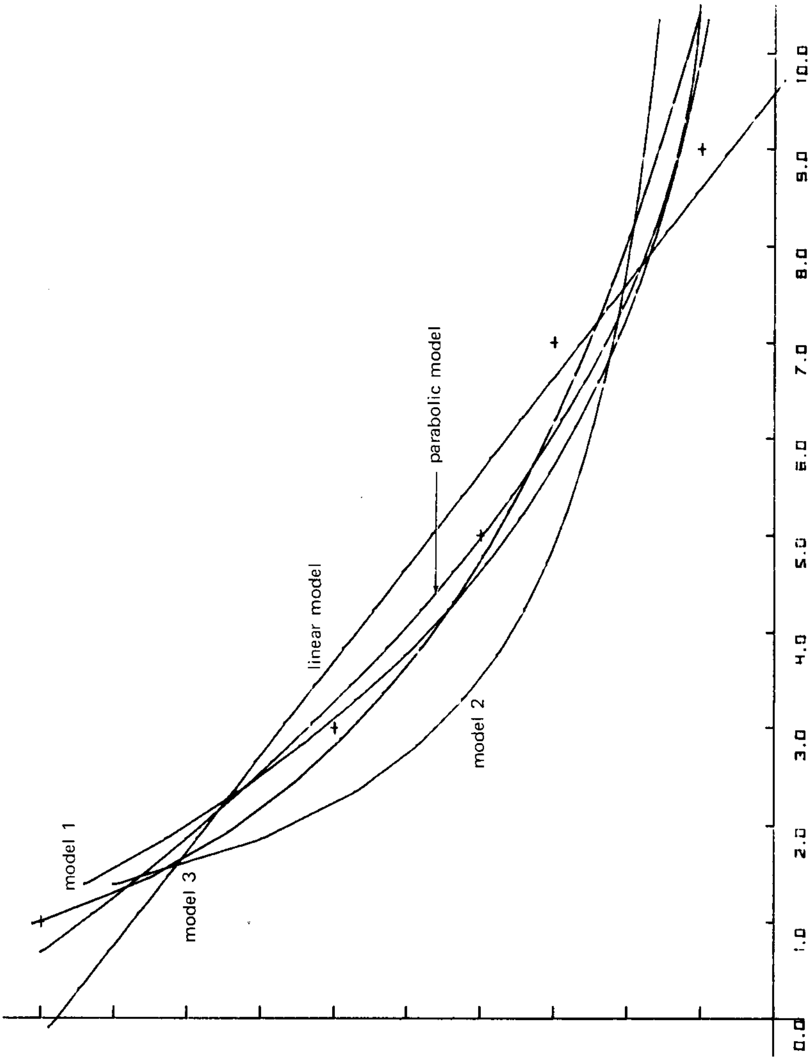
```
-        3.00      #
r²       A      B  =
          .99      #
        10.16      #
-        3.91      #

        MODL       =
```

11. Since all five models have been calculated, stop the program by pressing:

[LAST ENTRY] [RUN/STOP]

When the paper advances, the program is finished.

The next page shows an example of the axes, data points, and plotted lines generated by this example. The lines are numbered to correspond to the model which generated them.



Program 2: Family Regression (curve fitting)

To summarize:

linear r^2	= .94
parabolic r^2	= .98
r^2 for Model 1	= .94
r^2 for Model 2	= .79
r^2 for Model 3	= .99

Since model 3 has an r^2 of .99 you can conclude that 99% of the dependent variable's variation can be explained by using the logarithmic model.

Statistically, model 3 also proves to be the 'best fit' for this example (the best-fit, statistically, is found by using the r^2 value *and* the plot).

Program 3: One-Way Analysis Of Variance

This program performs a One-Way Analysis of Variance (AOV) on input data. Any number of treatment groups may be included, however each group must have at least two observations.

The One-Way AOV allows a test for statistical differences between the means of two or more groups which results in the F value being calculated.

Also, the One-Way AOV is a logical extension of the two-sample t-test. The two-sample t-test is a test for statistical difference between the means of two groups which results in a t value being calculated.

A table of t values and a table of F values are provided in Appendix A.

After the observations for each group are entered, the program calculates and prints these basic statistics:

N = number of observations in group

$\bar{x}$ = mean for group

s_1 = standard deviation for group

After all groups are entered and their basic statistics are calculated, these values are calculated and printed (see APPENDIX B for the formulas used):

$\bar{\bar{X}}$ = overall mean

SSB = Between Sum of Squares

V_1 = degrees of freedom for between treatments

MSB = Between Mean Squares

SSW = Within Sum of Squares

V_2 = degrees of freedom for within

MSW = Within Mean Square

V_1 = degrees of freedom for between

V_2 = degrees of freedom for within

F = F value

If the calculated F ratio is as large, or larger, than an F value obtained from a table of F values, then you can conclude that there is at least one group mean significantly different from at least one other group mean. To determine which groups differ significantly, a mean separation test such as Tukey's H.S.D. or Duncan's test should be used (see references in Appendix C). However, if the calculated F ratio is smaller than the appropriate tabled value, then you can conclude that there is no statistically significant difference between the group mean.

General Instructions

To perform a One-Way Analysis of Variance:

1. Start Program 3.
2. Key in each observation for the first group – press **DATA ENTRY** after keying in each observation.
3. When all observations are entered for the first group (you may use the **DELETE** key to remove incorrect entries), press **RUN/STOP**.

Typical Printout:

```
N                     =
        4.00          ¤  ←——  number of observations in group j
x̄                     =
        8.50          ¤  ←——  mean for group j
▲1                    =
        1.29          ¤  ←——  standard deviation for group j

V1   .......
```

The program is now waiting for you to enter the observations for the next group.

4. If there are more groups to be entered, return to step 2. After all groups have been entered and their respective basic statistics have been calculated, press **LAST ENTRY, RUN/STOP.**

Final Printout (typical results):

```
          GM          =
        7.58          ¤  ←——  overall mean

      231.92          ¤  ←——  Between Sum of Squares
        2.00          ¤  ←——  degrees of freedom from between
      115.96          ¤  ←——  Between mean Square

       23.00          ¤  ←——  Within Sum of Squares
        9.00          ¤  ←——  degrees of freedom from within
        2.56          ¤  ←——  Within Mean Square

V   1   2    F        =
        2.00          ¤  ←——  degrees of freedom from between
        9.00          ¤  ←——  degrees of freedom from within
       45.37          ¤  ←——  F value
```

After the program is finished the following data storage registers may be recalled.

Register Number	Contents
(constant storage)	MSW
0	ν_2
1	ν_1
2	$\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}^2$
3	MSB
4	$\sum_{j=1}^{k} n_j \bar{x}_j^2$

By using the final printout, a traditional analysis of variance table can be filled out as follows:

Source	df	SS	MS	F
Between	ν_1	SSB	MSB	F
Within	ν_2	SSW	MSW	—
Total	$\nu_1+\nu_2$	SSB+SSW	——	—

Example:

Run a One-Way AOV using these groups of observations:

Group 1	Group 2	Group 3
1	7	12
2	8	14
3	9	16
4	10	—
5	—	—

1. To start the program,

Press: [START PRGM]

CLEAR

and then **press** [3] [RUN STOP]

PRGM
3.00 *
DATACLEAR

2. To enter and calculate basic stat for the first group of observation,

Press: 1 (DATA ENTRY) 2 (DATA ENTRY) 3 (DATA ENTRY) 4 (DATA ENTRY) 5 (DATA ENTRY)

```
V1      •••••••
           1.00    #
           2.00    #
           3.00    #
           4.00    #
           5.00    #
```

Then press: (RUN STOP)

```
  N                =
           5.00    #
x̄                  =
           3.00    #
▲1                 =
           1.58    #
```

3. To enter and calculate basic stat for group 2,

Press: 6 (DATA ENTRY) 7 (DATA ENTRY) 8 (DATA ENTRY) 9 (DATA ENTRY) 10 (DATA ENTRY)

```
V1      •••••••
           6.00    #
           7.00    #
           8.00    #
           9.00    #
          10.00    #
```

To delete the incorrect entry (6),

Press: (DELETE)

and then press: 6 (DATA ENTRY)

```
              D E L
           6.00    #
```

Group 2 is now entered correctly,

so Press: (RUN STOP)

```
  N                =
           4.00    #
x̄                  =
           8.50    #
▲1                 =
           1.29    #
```

4. To enter and calculate basic stat for the last group,

Press: 12 [DATA ENTRY] 14 [DATA ENTRY]
16 [DATA ENTRY]

Then press: [RUN STOP]

```
V1      .......
          12.00    #
          14.00    #
          16.00    #

     N             =
           3.00    #

x̄                  =
          14.00    #

▲1                 =
           2.00    #
```

5. Since all groups are now entered, we can now calculate the AOV by pressing: [LAST ENTRY] [RUN STOP]

```
                        GM        =
x̄̄     ──────→         7.58       #

SSB   ──────→       231.92       #
ν1    ──────→         2.00       #
MSB   ──────→       115.96       #

SSW   ──────→        23.00       #
ν2    ──────→         9.00       #
MSW   ──────→         2.56       #

               V  1  2     F      =
ν1    ──────→         2.00       #
ν2    ──────→         9.00       #
F     ──────→        45.37       #
```

Now, traditional analysis of variance table can be filled out as follows:

Source	**df**	**SS**	**MS**	**F**
Between	2	231.92	115.96	45.37
Within	9	23.00	2.56	—
Total	11	254.92	—	—

From a table of F values (see Appendix A) notice that at the 95% level with 2 and 9 degrees of freedom the F value is 4.26. Since the calculated F value is 45.37, you can conclude that there is a significant difference between the means of the three groups.

Program 4: Normal Curve Overlay

This program is used to plot a normal curve based on the sample mean and standard deviation of histogram data in the calculator. Overlay plotting can occur either before or after a histogram has actually been plotted, provided that axes have been plotted and all histogram data has been entered into the calculator.

A normal curve overlay is useful in determining if a set of observations are normally distributed – a condition which is often assumed. If the data is normally distributed, then many statistical procedures are valid. If the data appears to deviate from the normality assumption, then some statistical test of normality may be indicated (see references in Appendix C for statistical tests of normality and alternate procedures in the event of non-normality).

After starting the program the basic statistics are calculated and printed and then the normal curve overlay is plotted. The formulas used to calculate basic stat in this program are listed in APPENDIX B.

General Instructions

1. Set up the plotter with paper and pen; then position the Graph Limits controls for the desired plot size.
2. Press [VAR #] [1] and then press [OFFSET WIDTH]

   Enter the offset width and press [DATA ENTRY]

   Enter the cell width and press [DATA ENTRY]
3. Now enter each observation (press [DATA ENTRY] after keying in each number).

NOTE

Although only observations between offset and (offset + 10 x cell width) are placed in histogram classes, *all observations* are counted when the normal curve overlay points are computed! Be sure to use the **DELETE** key to remove any unwanted observations before starting Program 4.

4. After all observations have been entered correctly,

   press [HISTO]

5. After the histogram is printed, set up the plotter axes by using Program 1 (X-axis limits are between 0 and the maximum cell number).

   Now, if you wish, you can press **PLOT** to plot the histogram.

6. Press [START PRGM] and then press [4] [RUN STOP]

   The normal curve overlay is now plotted. When the paper advances, the program is complete.

Example:

Plot a histogram and a normal curve overlay of the following list of observations.

1.00	2.00	4.00	8.00
2.00	3.00	5.00	5.00
3.00	4.00	6.00	6.00
4.00	5.00	7.00	7.00
5.00	6.00	8.00	6.00
6.00	7.00	4.00	7.00
7.00	8.00	5.00	6.00
8.00	9.00	6.00	
9.00	3.00	7.00	

To calculate the histogram:

1. **Press:** [VAR#] and **enter** 1	V1 •••••••
2. **Press:** [OFFSET/WIDTH] .5 [DATA ENTRY]	OF = .50 #
3. **Enter** 1 and **press** [DATA ENTRY]	C E L= 1.00 #
4. **Press** [DATA ENTRY] after entering each observation.	
5. After all observations are entered, **press** [HISTO]	

Printout:

```
                  HG                             6.00     #
cell number ──→ 1.00     #                       5.50     #
lower limit ──→  .50     #                       7.00     #
frequency ────→ 1.00     #                      20.59     #
% of n ───────→ 2.94     #
                                                 7.00     #
                2.00     #                       6.50     #
                1.50     #                       6.00     #
                2.00     #                      17.65     #
                5.88     #
                                                 8.00     #
                3.00     #                       7.50     #
                2.50     #                       4.00     #
                3.00     #                      11.76     #
                8.82     #
                                                 9.00     #
                4.00     #                       8.50     #
                3.50     #                       2.00     #
                4.00     #                       5.88     #
               11.76     #
                                                10.00     #
                5.00     #                       9.50     #
                4.50     #                        .00     #
                5.00     #                        .00     #
               14.71     #
```

6. Before plotting the histogram, start Program 1

by **pressing:** [START PRGM]

CLEAR

Then **press:** [1] [RUN STOP]

PRGM
1.00 #

To determine the scale limits, look at the histogram printout to find the last cell number with observations (cell 9). In this case: X_{max} = 9, X_{min} = 0, X_{tic} = 1. Also, the highest cell frequency is 7; therefore; Y_{max} = 7, Y_{min} = 0, Y_{tic}= 1.

7. To enter the plot dimensions and plot the axes,

Press: 0 [RUN STOP] 9 [RUN STOP] 0 [RUN STOP]

Press: 7 [RUN STOP] 1 [RUN STOP] 1 [RUN STOP]

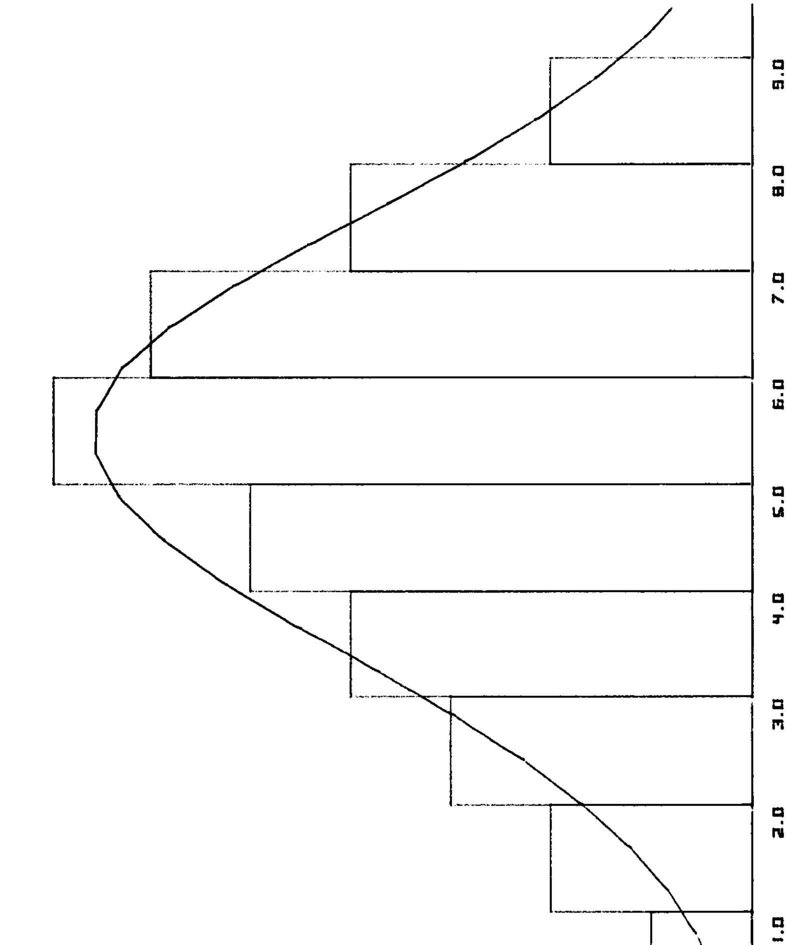
```
    .00   #
   9.00   #
    .00   #
   7.00   #
   1.00   #
   1.00   #
     AI   =
    .00   #
    .00   #
```

After the program stops,

Press: [RUN STOP]

After the program stops again, **press:** [RUN STOP]

8. After the axes are plotted and labeled, plot the histogram by **pressing** [PLOT]



Program 4: Normal Curve Overlay

9. To plot the normal curve overlay,

Press: [START PRGM]

```
                    CLEAR
```

Then press: [4] [RUN STOP]

```
              PRGM
               4.00     #
    N                   =
              34.00     #
 x̄                      =
               5.56     #
 ▲1                     =
               2.05     #
```

By looking at the normal curve overlay superimposed upon the histogram, it appears that the data for this example does not deviate excessively from normality. The histogram shows a skew to the left but it doesn't appear to be significant.

Expanded Data Storage

When the Expanded Stat Programs were installed in your calculator, an extra 30 data registers were also installed (each register is a place for storing one number). Although the extra data storage is in use when a program is running, you may use each register just as you would use constant storage when doing keyboard calculations.

The extra data registers are numbered 0 through 29.

Store And Recall

To store the current number into an extra storage register, first enter the number, then press:

[SHIFT] [STORE] [REGISTER NUMBER] [=]

Later, to recall the number from an extra storage register, press:

[SHIFT] [RECALL] [REGISTER NUMBER] [=]

If the calculator prints 'NOTE 26' after you tried to store into, or recall from, a data register, the register number you specified is not in the calculator.

Listing Extra Storage Registers

When you wish to list the numbers in all the extra storage registers, switch the printer on and press:

[SHIFT] [START PRGM] [PRINT] [=]

The data is listed in blocks of ten, beginning at data register number '0'.

Here's what a data register listing looks like:

```
          DATA     =

        100.00          <------ register no. 0
        100.00
         25.00
-        45.00
           .71
          3.14
-        62.00
         62.70
           .00
9.999999999   99        <------ register no. 9

           .00          <------ register no. 10
           .00

-          .00
           .01
         40.79
           .00
          6.92
-        15.00          <------ last register (29)

              END
```

Clearing Extra Storage Registers

To erase the data from all extra storage registers in your calculator, press:

SHIFT START PROGM 0 =

'DATACLEAR' is printed after the storage registers are erased.

Chapter **3**

Sample Applications

This chapter describes thirteen problems that were solved by using the Statistics Calculator and the Expanded Stat Programs. **Here's a list of the applications:**

Business and Management (page 32):

- Family Regression – product sales analysis
- One-Way AOV – product evaluation

Industry (page 35):

- Family Regression – effects of chemical dilutions
- One-Way AOV – product evaluation
- Histogram – quality control

Research and Development (page 41):

- Family Regression – engine performance
- One-Way AOV – comparison of fuel additives
- Histogram – life-testing new components

Biological Sciences (page 48):

- Linear Regression – genetics study
- One-Way AOV – effects of animals' diets
- Histograms:

  A – Birth rate
  B – Effects of drug dosages

The solution to each application is described in an abbreviation form (all steps are not listed) and assumes that you are familiar with operating the calculator and the Expanded Stat Programs.

Business And Management

Family Regression

After introducing their newest product, the ABC company wanted to look at the percentage of increase in sales as a function of the number of months since the product was introduced. The following data were obtained:

x: months since introduction	1	2	3	4	5	6
y: percent increases in sales	3	6	15	30	60	95

Since the x values range between 1 and 6 in integer values, the X-axis dimensions are:

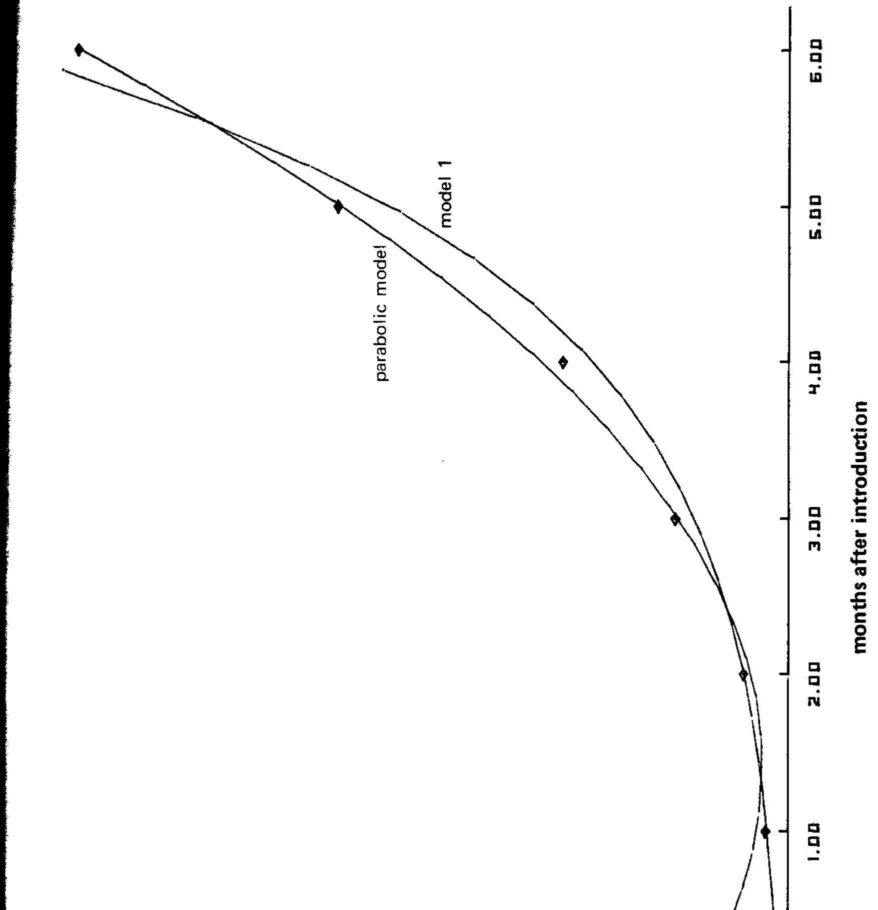
$$X_{min} = 0$$
$$X_{max} = 6$$
$$X_{tic} = 1$$

Since the y values range from 3 to 95, the Y-axis dimensions are:

$$Y_{min} = 0$$
$$Y_{max} = 95$$
$$Y_{tic} = 10$$

After setting up the plotter to those limits, we entered the data. Character 2 was selected for plotting. The finished plot is on the following page.

On the basis of r^2, the parabolic model produces the "best" fit. However, A plot of the predicted line reveals a marked departure from theory around the origin. Model 1 (exponential) obtained an r^2 value of .99, and a plot of the fitted line *does* agree with the theoretical results. Since growth starts at or near zero and then increases, we consider the exponential model (model 1 to be the "best" equation for this set of data.


% increase in sales
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
months after introduction
0.00
1.00
2.00
3.00
4.00
5.00
6.00
parabolic model
model 1

One-Way AOV

An electronics plant manager thinks that there is a difference in the rates of absenteeism among three groups of employees. After checking with the payroll office he comes up with the following figures:

(Days Missed In Six Months)

Engineers	Assembly	Sales Support
2	13	5
2	12	3
3	10	7
5	15	6
0	13	5
3	14	
	10	

A one-way analysis of variance reveals the following table:

Source	df	SS	MS	F
Between Groups	2	343.76	171.88	58.58
Within Groups	15	44.01	2.93	––

The 'tabled' F value at the 95% level with $\nu_1=2$ and $\nu_2=15$ is 3.68. Since our calculated F value is 58.60, we can conclude that there is a significant difference between the number of days missed by the three groups of employees.

Printout:

```
V1     .......

          GM       =
         7.11      #

       343.76      #
         2.00      #
       171.88      #

        44.01      #
        15.00      #
         2.93      #

V  1  2      F     =
         2.00      #
        15.00      #
        58.58      #
```

Industry

Family Regression

A quality control engineer notes that there seems to be a relationship between the amount of a chemical added to a batch of a liquid product and the final concentration of the chemical in the final product. The following data was obtained:

weight added (x):	2	1	6	3	7	6	9
wt. of final product (y):	3	1	5	5	7	8	8.5

Since the x values range from 1 to 9, the X-axis dimensions are:

$$X_{min} = 0$$
$$X_{max} = 9$$
$$X_{tic} = 1$$

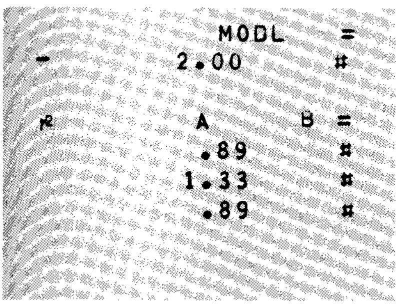
Since the y values range from 1 to 8.5, the Y-axis dimensions are:

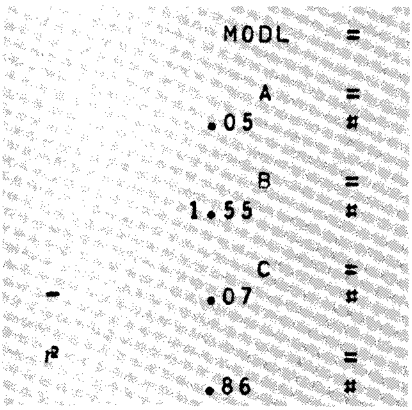
$$Y_{min} = 0$$
$$Y_{max} = 8.5$$
$$Y_{tic} = 1$$

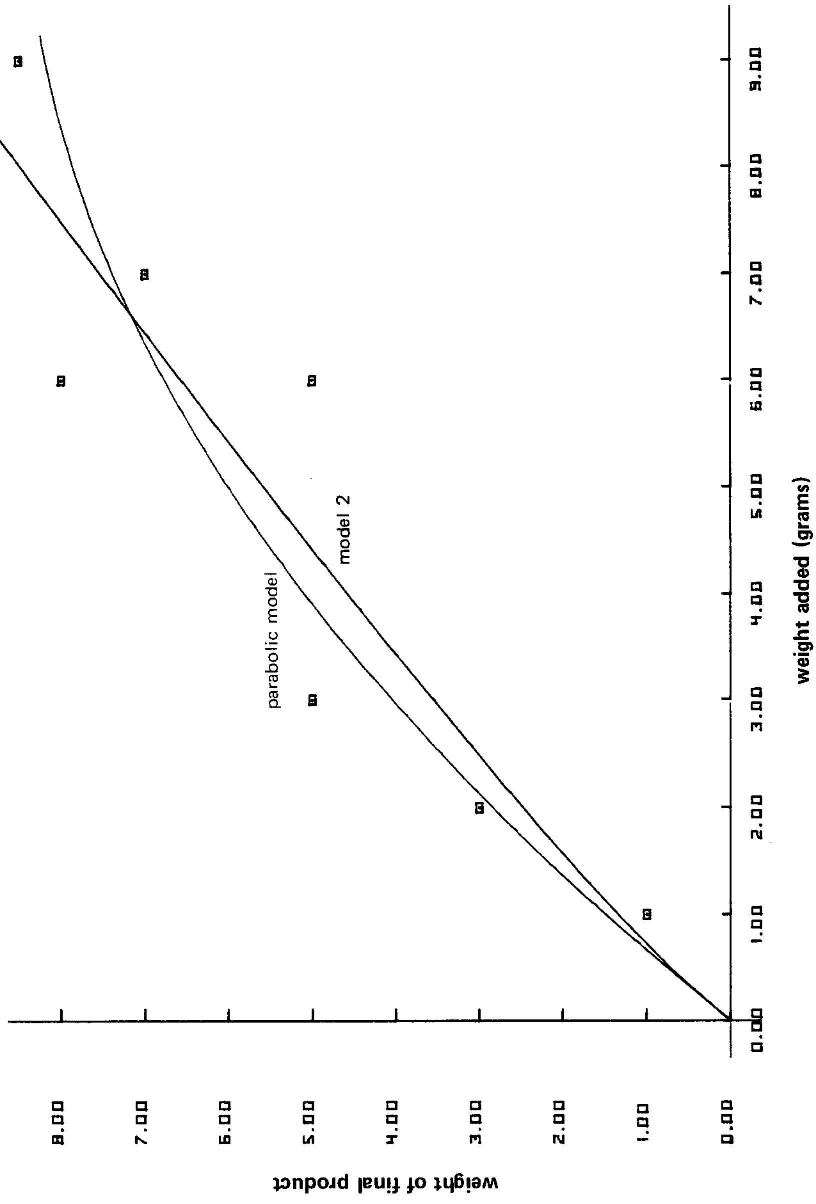
After setting the plotter to those limits, the data is entered. Character 3 was used to plot the data (see the plot on the next page).

On the basis of squared correlation coefficient, Model 2 (the power curve) is judged to be the "best fit" for this set of data.

Final Printout:







weight of final product
weight added (grams)
parabolic model
model 2
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00

One-Way AOV

Three brands of typewriters were compared as to key pressure necessary to produce a "clean" impression when the pressure indicator was set at "Light". After four machines of each brand were randomly selected these measurments were obtained:

(Pressure in Grams)

Brand A	Brand B	Brand C
100	150	250
125	155	310
90	130	290
110	140	275

A one-way analysis of variance yields the following table (see the following printout):

Source	df	SS	MS	F
Between	2	67916.67	33958.33	103.38
Within	9	2956.25	328.47	—

The 'tabled' F value at the 95% level with $\nu_1=2$, $\nu_2=9$ is 4.26. Since the calculated F value is 103.38, the hypothesis of no difference between the brands of typewriters can be rejected. By looking at the individual typewriter means, notice that Brand A seems to have the lightest touch.

Printout:

```
        Brand A

V1      .......
        100.00     #
        125.00     #
         90.00     #
        110.00     #

  N                =
          4.00     #

x̄                  =
        106.25     #

▲1                 =
         14.93     #

        Brand B

V1      .......
        150.00     #
        155.00     #
        130.00     #
        140.00     #

  N                =
          4.00     #

x̄                  =
        143.75     #

▲1                 =
         11.09     #
```

Brand C

```
V1       .......
          250.00     #
          310.00     #
          290.00     #
          275.00     #

   N                 =
            4.00     #

x̄                    =
          281.25     #

▲1                   =
           25.29     #
```

```
              GM     =
          177.08     #

        67916.67     #
            2.00     #
        33958.33     #

         2956.25     #
            9.00     #
          328.47     #

V  1  2        F     =
            2.00     #
            9.00     #
          103.38     #
```

Histogram

Samples of castings for a certain machine part are inspected at the end of a week's production. The following data represent the observed number of imperfections for each unit. The production manager is interested in looking at the distribution of these defects each week.

Unit Number	Number of Defects	Unit Number	Number of Defects
1	0	16	0
2	8	17	6
3	1	18	3
4	7	19	9
5	3	20	1
6	9	21	8
7	2	22	0
8	9	23	7
9	0	24	0
10	8	25	9
11	3	26	2
12	6	27	9
13	0	28	1
14	1	29	8
15	2	30	0

Since the data range from 0 to 9 defects,

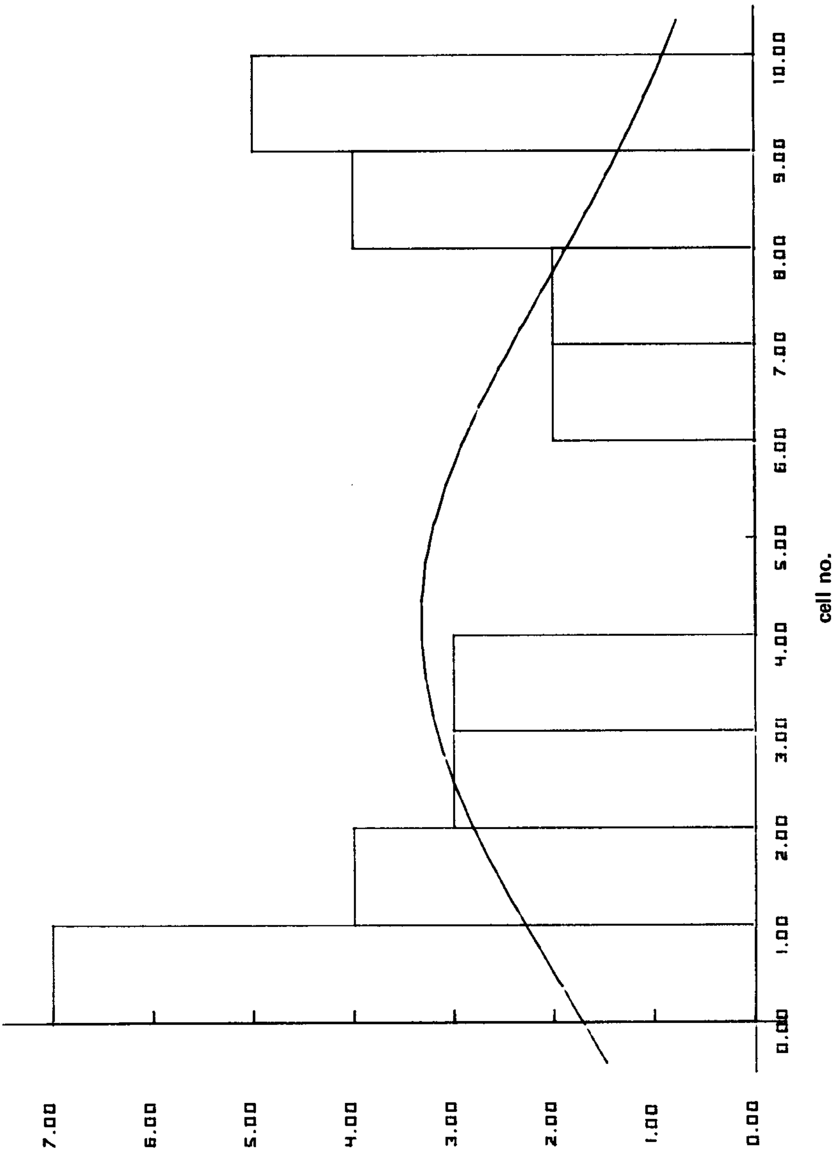
Offset = 0
Cell Width = 1

After entering the data into the calculator, the following histogram printout was obtained.

```
                    HG
cell number ---->  1.00   #          6.00   #
lower limit ---->   .00   #          5.00   #
frequency   ---->  7.00   #           .00   #
% of n      ----> 23.33   #           .00   #

                   2.00   #          7.00   #
                   1.00   #          6.00   #
                   4.00   #          2.00   #
                  13.33   #          6.67   #

                   3.00   #          8.00   #
                   2.00   #          7.00   #
                   3.00   #          2.00   #
                  10.00   #          6.67   #

                   4.00   #          9.00   #
                   3.00   #          8.00   #
                   3.00   #          4.00   #
                  10.00   #         13.33   #

                   5.00   #         10.00   #
                   4.00   #          9.00   #
                    .00   #          5.00   #
                    .00   #         16.67   #
```

By looking at the printout, note that there are observations in cells 1 and 10, and the highest frequency is 7. From those values, the axis deminsions should be:

$X_{min} = 0$
$X_{max} = 10$
$Y_{min} = 0$
$Y_{max} = 7$
$X_{tic} = 1$
$Y_{tic} = 1$


no. of defects
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
cell no.
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00
10.00

After the histogram and overlay were plotted, notice that the histogram appears to be 'U' shaped rather than normal. There appears to be two distinct groups, which indicates that perhaps we should reexamine the way the observations were taken:

- The histogram could indicate that one or more of the casting machines is seriously malfunctioning.
- We could also look for possible effects due to the day of the week that the item was produced. If production were under control, we might expect a histogram similar to that obtained for classes 1 through 4.

Research And Developement

Family Regression

The thrust of a certain jet engine seems to be related to the size of the exhaust orifice. Some tests were run to establish that relationship; they yielded the following observations (x is the exhaust orifice size in centimeters, y is the attained thrust):

x:	.001	.004	.002	.006	.009
y:	.014	.007	.010	.002	.001

Since the range of the x variable is .00l to .009, the X-axis dimensions are:

$$X_{min} = .000$$
$$X_{max} = .009$$
$$X_{tic} = .001$$

Since the range of the y variable is .001 to .014, the Y-axis dimensions are:

$$Y_{min} = .000$$
$$Y_{max} = .014$$
$$Y_{tic} = .002$$

> **NOTE**
>
> Since the **AUTO·** feature cannot be used while running these programs, be sure the [AUTO·] key is up (switched off)!

While using program 1 to set up and label the axes, set the printing format for labeling the X_{tic} marks by pressing:

[SHIFT] [PRINT] [1] [RND( )] 3 [LAST ENTRY] [RUN/STOP]

After the display returns for setting the printing format for labeling the Y_{tic} marks, press the above sequence again.

After setting the plotter to those limits the data was entered according to the program's instructions. Character 0 was used to plot the data points.

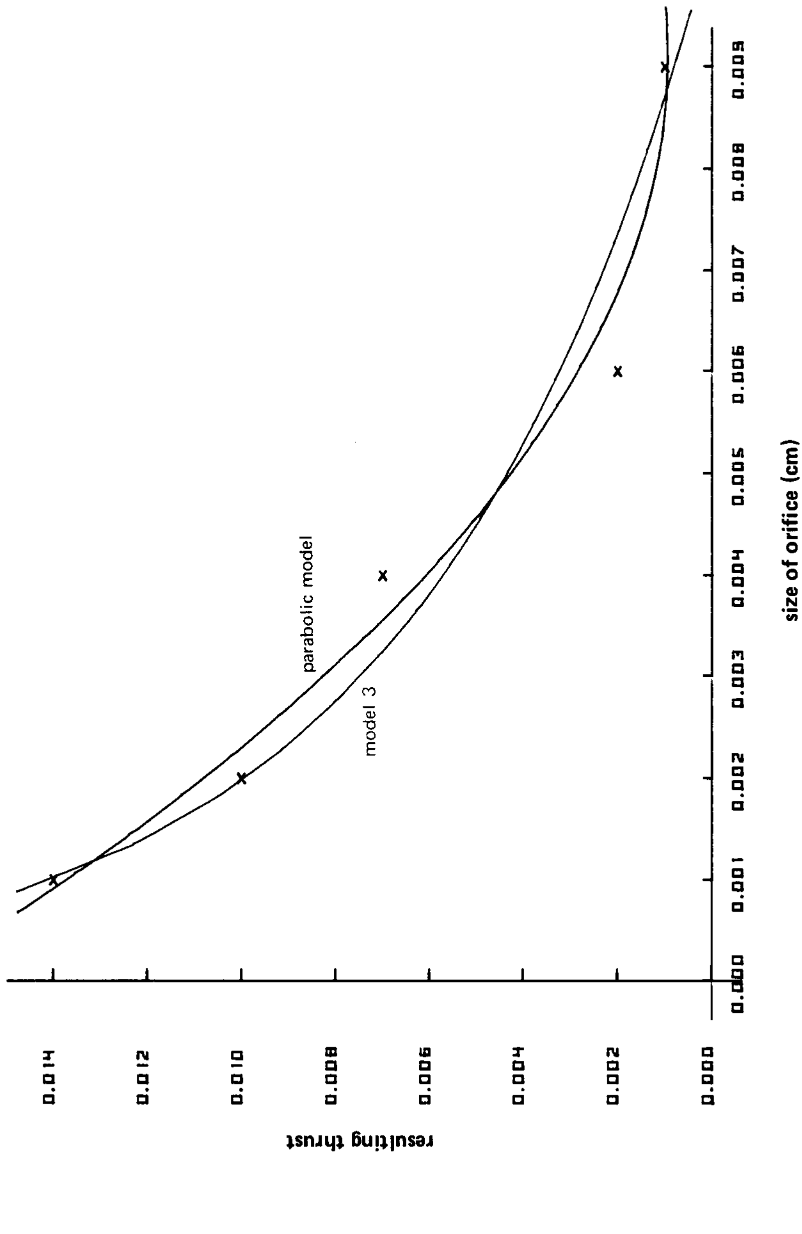
On the basis of squared correlation coefficient, the parabolic model was judged as the 'best fit' for this set of data.

Parabolic

```
            A  =
         .017  #

            B  =
-       3.594  #

            C  =
      198.344  #

r²             =
         .981  #
```

Logarithmic

```
       MODL    =
-     3.000    #

r²     A     B =
      .972     #
-     .028     #
-     .006     #
```


resulting thrust
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0.009
size of orifice (cm)
parabolic model
model 3

One-Way AOV

A production engineer wants to compare three new types of gasoline additives with the company's standard gasoline additives. He obtained a random sample of three observations on the performance of each additive, and generated the following data (expressed in miles per gallon):

Additive A	Additive B	Additive C	Standard Additive
10.1	15.2	11.2	8.7
7.2	19.1	12.1	10.8
12.3	13.8	10.9	13.6

A one-way analysis of variance of the four groups yields the following table (also see the following printout):

Source	df	SS	MS	F
Between Groups	3	66.26	22.09	4.31
Within Groups	8	41.04	5.13	––

The 'tabled' F value at the 95% level with $\nu_1=3$ and $\nu_2=8$ is 4.07. Since the calculated F value is 4.31, the hypothesis of no difference between the four additives is rejected. By looking at each groups' mean, notice that Additive B seems to produce the best gasoline mileage.

Printout:

```
        GM        =
     12.08        #

     66.26        #
      3.00        #
     22.09        #

     41.04        #
      8.00        #
      5.13        #

V 1 2      F      =
      3.00        #
      8.00        #
      4.31        #
```

Histogram

In developing performance standards for a new type of electrical device, 16 items are tested. The data below indicate hours-to-failure of the sixteen items.

item	hours to failure	item	hours to failure
1	2.9	9	0.5
2	1.0	10	2.3
3	3.2	11	4.8
4	1.5	12	1.9
5	1.0	13	7.3
6	5.0	14	2.8
7	0.7	15	11.1
8	6.5	16	15.7

The range of the data is .5 to 15.7 If the offset is set to 0 and the cell width is set to 2, the histogram will contain 8 cells.

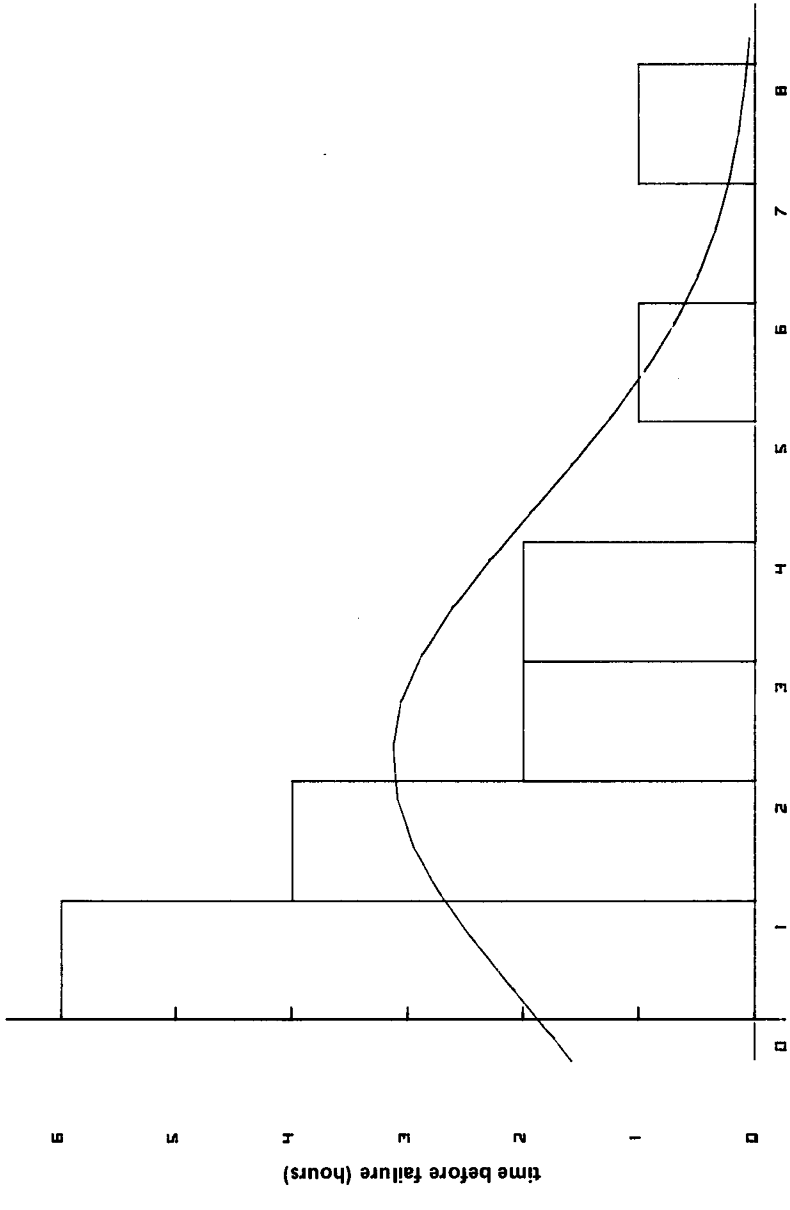
After the data has been entered and a histogram generated, notice that cell 8 is indeed the last cell containing any observations. Also, note that the highest frequency is 8. Therefore, the axes dimensions can be specified as follows:

$$X_{min} = 0$$
$$X_{max} = 8$$
$$Y_{min} = 0$$
$$Y_{max} = 6$$
$$X_{tic} = 1$$
$$Y_{tic} = 1$$

The output and plot contained on the following pages indicate a highly skewed distribution. Since electrical components are being observed, experience dictates that this type of skewed distribution can be expected. Also, notice that the normal curve overlay is relatively flat and does not coincide with the mode, or high point, of the histogram. These results indicate that we need to look at other kinds of distributions than normal for this set of data.

```
     HG
   1.00    #
    .00    #
   6.00    #
  37.50    #

   2.00    #
   2.00    #
   4.00    #
  25.00    #

   3.00    #
   4.00    #
   2.00    #
  12.50    #

   4.00    #
   6.00    #
   2.00    #
  12.50    #

   5.00    #
   8.00    #
    .00    #
    .00    #

   6.00    #
  10.00    #
   1.00    #
   6.25    #

   7.00    #
  12.00    #
    .00    #
    .00    #

   8.00    #
  14.00    #
   1.00    #
   6.25    #

   9.00    #
  16.00    #
    .00    #
    .00    #

  10.00    #
  18.00    #
    .00    #
    .00    #
```

```
                 CLEAR



            PRGM
             4.00     #


      N               =
            16.00     #

x̄                     =
             4.26     #

▲1                    =
             4.19     #
```


cell no.
time before failure (hours)

Biological Sciences

Linear Regression

The following data represent a hypothetical genetic study (x values refer to generation number and the y values are weights, in grams, of a bacterial colony).

i	1	2	3	4	5	6	7	8
x_i	1	1	2	2	3	3	4	4
y_i	2.72	1.49	3.32	5.47	16.45	25.21	66.69	81.45

Notice that the x values are the integers 1, 2, 3, 4. To make a usable graph, a margin should be allowed on both sides with tri-intervals at 1, 2, 3 and 4. Thus, specify these X-axis dimensions (use Program 1):

$$X_{min} = 0$$
$$X_{max} = 4$$
$$X_{tic} = 1$$

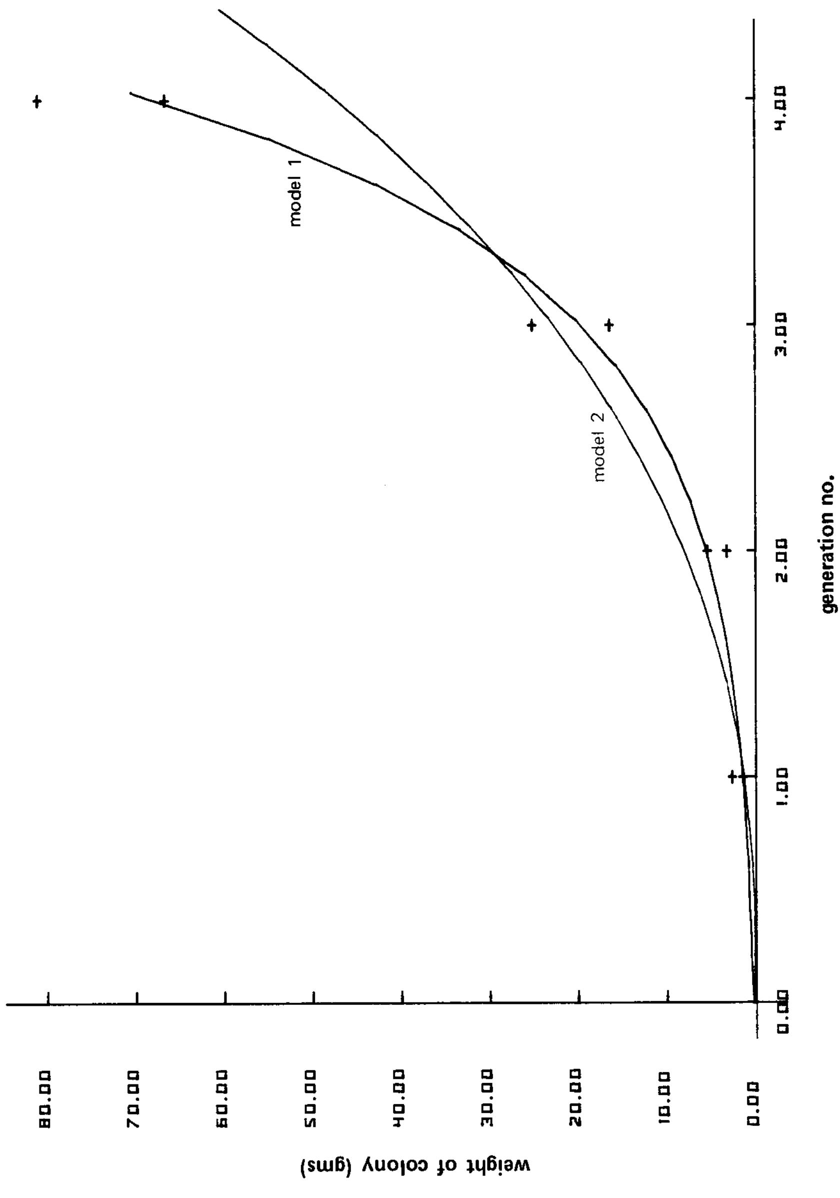
Our y values range from 2.72 up to 81.45. To allow a margin at both top and bottom, specify these Y-axis dimensions:

$$Y_{min} = 0$$
$$Y_{max} = 85$$
$$Y_{tic} = 10$$

After setting up the axes we're ready to perform the regression analysis using Program 2.

Once the data was entered and plotted, an examination of the plotted points revealed a curvi-linear relationship. Since bacteria growth (and hopefully weight) increases exponentially, we can expect that either model 1 or 2 may provide the 'best fit'. The following printout indicates that model 1 (exponential) provides the best fit, as measured by the value of the squared correlation coefficient (r^2 = .96).

```
            MODL      =                     MODL      =
  -         1.00      ¤          -          2.00      ¤

  r²        A      B  =          r²         A      B  =
            .96       ¤                     .88       ¤
            .48       ¤                    1.41       ¤
           1.24       ¤                    2.54       ¤
```



model 1
model 2
generation no.
weight of colony (gms)
0.00
1.00
2.00
3.00
4.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00

One-Way AOV

The following data represent weight gains in experimental animals fed on three different diets.

(Weight Gains are Expressed in Grams)

Diet A	Diet B	Diet C
15	13	15
20	17	18
19	22	12
*	15	21

A one-way analysis of variance yields the following (also see the printout):

Source	df	SS	MS	F
Between	2	4.25	2.13	.16
Within	8	103.75	12.97	––

The 'tabled' F value at the 95% level with $\nu_1=2$, $\nu_2=8$ is 4.46. Since our calculated F value of .16 is considerably less than 4.46 we can conclude that, based on our observations, there's no statistical evidence of a significant difference between the three diets.

Printout:

```
          GM      =
       17.00      #

        4.25      #
        2.00      #
        2.13      #

      103.75      #
        8.00      #
       12.97      #

V  1  2    F      =
        2.00      #
        8.00      #
         .16      #
```

*Animal 4 of Group A died of heart disease.

Histogram A

Many people believe that time of birth is distributed normally about a mean of midnight. The following table represents birth time in hours-from-noon for non-induced births during a given period at a small-town hospital.

7	11	16	20	21	12
13	20	2	12	22	2
6	21	4	20	15	23
3	18	5	1	2	12
17	18	19	13	1	13
3	4	8	2	22	23
5					

The data ranged from 0 = 12:00 a.m. (noon) to 24 which is again 12:00 noon. The histogram function on the calculator allows a maximum of 10 cells, but if 8 cells (each cell width is 3) are used, the range will be covered.

First, set up the one-variable entry mode, and let:

Offset = 0
Cell Width = 3

Then enter the data using the **DATA ENTRY** key. Next, calculate the histogram and observe that indeed only 8 cells have frequencies! Also, note that the highest frequency is 6. Therefore, specify the following plotter dimensions:

$X_{min} = 0$
$X_{max} = 8$
$Y_{min} = 0$
$Y_{max} = 6$
$X_{tic} = 1$
$Y_{tic} = 1$

By looking at the printout and plot, notice that the histogram indicates a wide departure from the "bell shaped" normal curve. It appears that a rectangular distribution might be a better description of the observed birth times. Note that the X-axis labeling refers to cell number – *not* actual class boundaries.

HG

1.00 #
.00 #
6.00 #
16.22 #

2.00 #
3.00 #
6.00 #
16.22 #

3.00 #
6.00 #
3.00 #
8.11 #

4.00 #
9.00 #
1.00 #
2.70 #

5.00 #
12.00 #
6.00 #
16.22 #

6.00 #
15.00 #
3.00 #
8.11 #

7.00 #
18.00 #
6.00 #
16.22 #

8.00 #
21.00 #
6.00 #
16.22 #

9.00 #
24.00 #
.00 #
.00 #

10.00 #
27.00 #
.00 #
.00 #

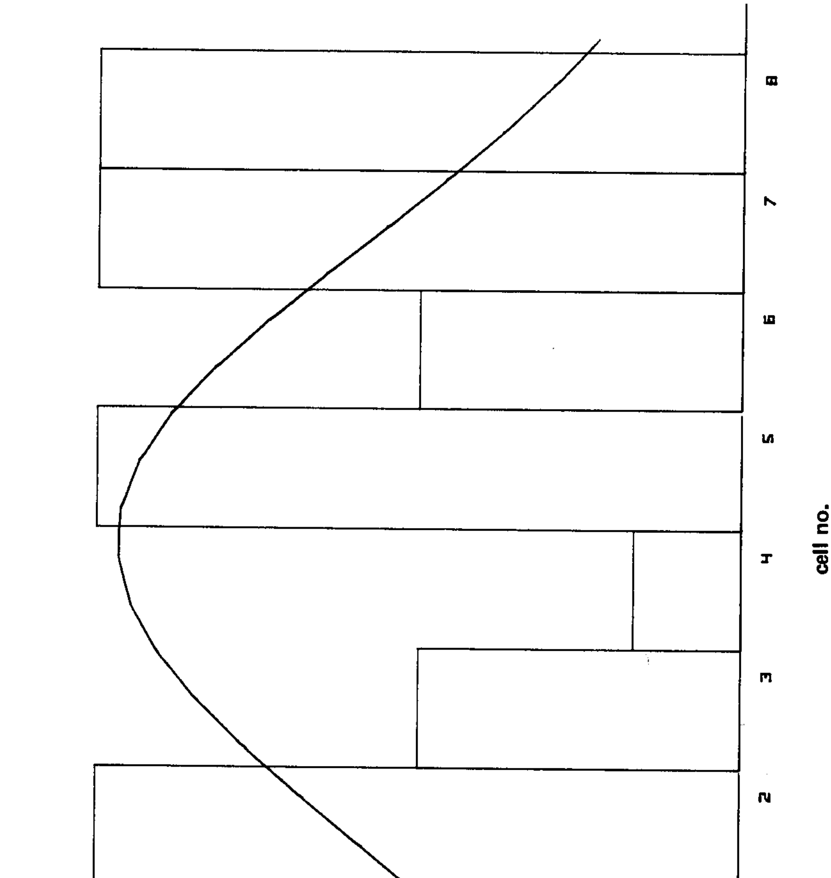
CLEAR

PRGM
4.00 #

N =
37.00 #

$\bar{x}$ =
11.78 #

Δ1 =
7.59 #


birth time (hours-to-noon)
cell no.
0
1
2
3
4
5
6
7
8

Histogram B

An experiment was run to measure fluid retention of subjects treated with a new drug.

Two groups of subjects were administered tablets. Group one received the drug while group two received a sugar pill (placebo). Their measured fluid retention (expressed in cc of liquid) are given below.

Group 1

Subject No.	cc Retained
1	26
2	41
3	47
4	44
5	41
6	49
7	43
8	40
9	48
10	37
11	44
12	39
13	42
14	36
15	33

Group 2

Subject No.	cc Retained
1	1
2	4
3	3
4	6
5	3
6	9
7	12
8	4
9	18
10	8
11	2
12	7
13	13

The experimenter is interested in determining the difference, if any, between the two groups. By running a two-sample t-test, the following printout was obtained:

```
N1               =
     15.00       #
x̄                =
     40.67       #
▲1               =
      6.01       #

N2               =
     13.00       #
ȳ                =
      6.92       #
▲2               =
      4.99       #

T                =
     16.01       #
     D           =
     26.00       #
```

The 'tabled' t value at the 95% level with 26 degrees of freedom is 2.0567. Since our calculated t value was found to be 16.01, we can reject the hypothesis of no difference between the means of the two groups.

Since our t-test shows a significant difference between the groups, a graphical presentation of this difference is desired. Note that the range of the two samples is from 1 to 49. If we wished to get a histogram of the total data, an offset of 0 and a cell width of 5 would give 10 classes of width 5 each. After entering the cell width and offset, enter the data for Group 1 and plot the resulting histogram and normal curve overlay. Then reset the calculator for generating Group 2's histogram and normal curve overlay. Be certain to press **AXES, 0,** to preserve the previous plot dimensions. The resulting printout and graph on the following pages clearly illustrates the difference between the two groups.

We should notice that Group 1 is represented by cells 6 through 10, while Group 2 is represented by cells 1 through 4. Although the histogram for Group 1 is close to the 'bell shape', the histogram for Group 2 seems to be quite different from the bell shape. In view of the extremely large t value, this slight departure may be negligible.

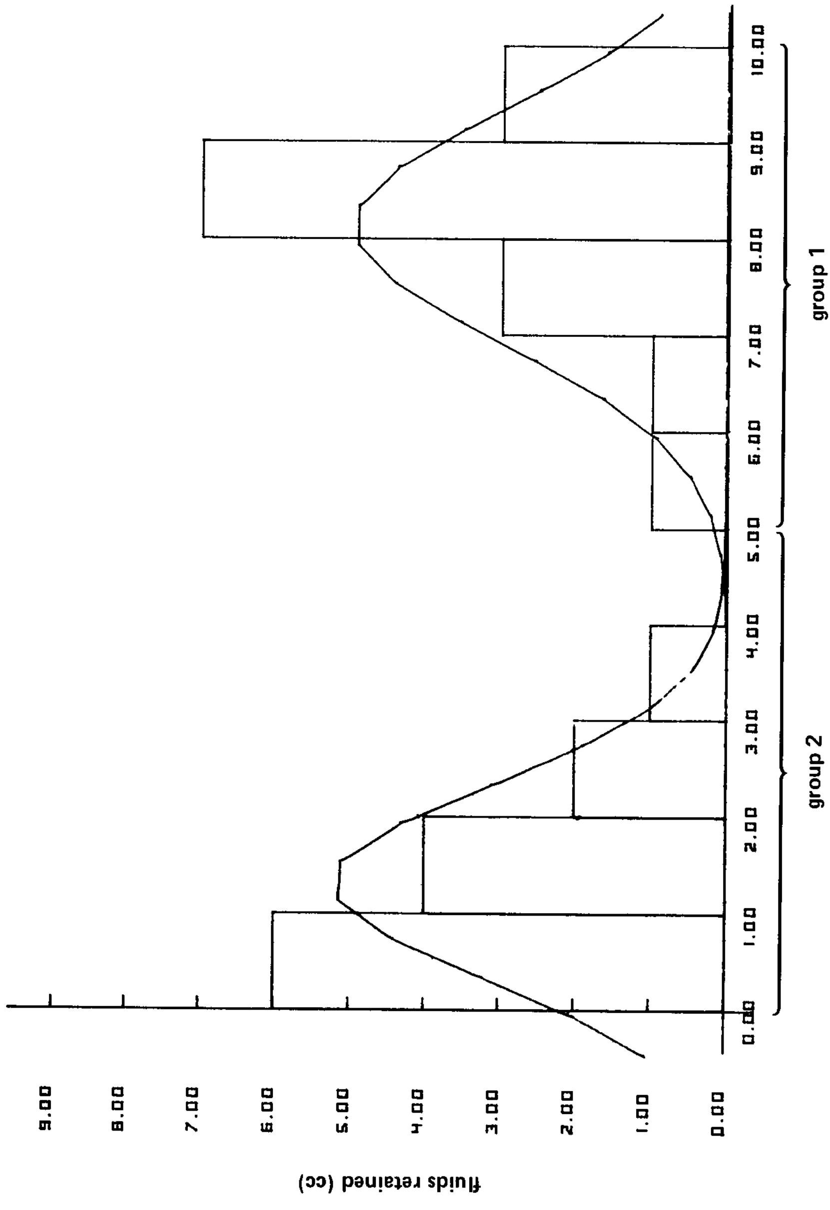
We can conclude that, based on the t statistic and the plotted histograms, there appears to be a very dramatic difference between these two groups.

Group 1

HG	
1.00	#
.00	#
.00	#
.00	#
2.00	#
5.00	#
.00	#
.00	#
3.00	#
10.00	#
.00	#
.00	#
4.00	#
15.00	#
.00	#
.00	#
5.00	#
20.00	#
.00	#
.00	#
6.00	#
25.00	#
1.00	#
6.67	#
7.00	#
30.00	#
1.00	#
6.67	#
8.00	#
35.00	#
3.00	#
20.00	#
9.00	#
40.00	#
7.00	#
46.67	#
10.00	#
45.00	#
3.00	#
20.00	#

Group 2

HG	
1.00	#
.00	#
6.00	#
46.15	#
2.00	#
5.00	#
4.00	#
30.77	#
3.00	#
10.00	#
2.00	#
15.38	#
4.00	#
15.00	#
1.00	#
7.69	#
5.00	#
20.00	#
.00	#
.00	#
6.00	#
25.00	#
.00	#
.00	#
7.00	#
30.00	#
.00	#
.00	#
8.00	#
35.00	#
.00	#
.00	#
9.00	#
40.00	#
.00	#
.00	#
10.00	#
45.00	#
.00	#
.00	#


fluids retained (cc)
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00
10.00
group 2
group 1

Notes

Appendix **A**

Tables

Percentage Points of the t-Distribution

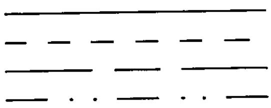
ν	90%	95%	98%	99%
1	6.31	12.7	31.8	63.7
2	2.92	4.31	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.90	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
12	1.78	2.18	2.68	3.06
14	1.76	2.14	2.62	2.98
16	1.75	2.12	2.58	2.92
18	1.73	2.10	2.55	2.88
20	1.72	2.09	2.53	2.84

Upper 95% Confidence Level for the F-Distribution

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35

Plotting Formats *

Plotting Code	Description	Example Plot
0	solid line (automatically set)	
1	dashed line	
2	alternating long and short dashes	
3	alternating dashes and dots	

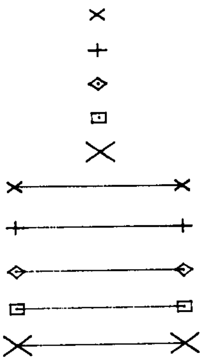


To specify (or change) the plotting format,

press: [SHIFT] [PRINT] [1] [)] [PLOT CODE] [LAST ENTRY]

Plotting Characters *

Character Code	Example Plot
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	



To specify the plotting character, **press:** [CHAR #] [CHARACTER CODE]

*The plotting code and plotting character specified will remain unchanged until either they are changed or the calculator is switched off. A solid-line format (code 0) and the plotting character 'X' (code 0) are automatically set when the calculator is switched on.

Appendix **B**

Formulas

Program 2: Family Regression

Notation:

y_i = Dependent or response variable

x_i = Independent or predictor variable

n = Number of (x_i, y_i) pairs, i=1, 2, . . ., n

$\bar{x}$ = Mean of the independent variable

$\bar{y}$ = Mean of the dependent variable

$$U_x = \sum_{i=1}^{n} (x - \bar{x})^2$$

$$U_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

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

$$U_{xx}{}^2 = \sum_{i=1}^{n} (x_i - \bar{x})(x_i^2 - \overline{x^2})$$

$$U_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$U_{x^2y} = \sum_{i=1}^{n} (x_i^2 - \overline{x^2})(y_i - \bar{y})$$

$$U_{\ln x} = (\ln x_i - \frac{1}{n} \sum_{n=1}^{n} \ln x_i)$$

$$U_{\ln y} = (\ln y_i - \frac{1}{n} \sum_{i=1}^{n} \ln y_i)$$

Linear Regression:

a = Intercept =

$$\frac{1}{n}[\Sigma y - b\Sigma x]$$

b = Slope =

$$\frac{U_{xy}}{x}$$

r^2 = Square of the correlation coefficient =

$$\frac{U_{xy}{}^2}{U_x U_y}$$

Parabolic Regression:

a = Intercept =

$$\frac{1}{n}[\Sigma y - cx^2 - b\Sigma x]$$

b = Linear coefficient =

$$\frac{U_{xy} - cU_{xx^2}}{U_x}$$

c = Quadratic coefficient =

$$\frac{U_x U_{x^2 y} - U_{xx^2} U_{xy}}{U_x U_{x^2} - (U_{xx^2})^2}$$

r^2 = Square of the multiple correlation coefficient =

$$\frac{bU_{xy} + cU_{x^2 y}}{U_y}$$

Exponential Curve:

a = Constant =

$$\exp \frac{1}{n} \sum_{i=1}^{n} \ln y_i - bx$$

b = x coefficient =

$$\frac{\sum_{i=1}^{n} (x_1 - \bar{x})\, U_{\ln x}}{U_x}$$

r^2 = Square of the correlation coefficient =

$$\frac{\sum_{i=1}^{n} (x_1 - \bar{x})\, U_{\ln y}}{U_x \sum_{i=1}^{n} (U_{\ln y})^2}$$

Power Curve:

a = Constant =

$$\frac{1}{n} \sum_{i=1}^{n} \ln y_i - \frac{b}{n} \sum_{i=1}^{n} \ln x_i$$

b = Power =

$$\frac{\sum_{i=1}^{n} U_{\ln x} U_{\ln y}}{\sum_{i=1}^{n} (U_{\ln x})^2}$$

r^2 = Square of the correlation coefficient =

$$\frac{\sum_{i=1}^{n} U_{\ln x} (U_{\ln y})^2}{\sum_{i=1}^{n} (U_{\ln x})^2 \sum_{i=1}^{n} (U_{\ln y})^2}$$

Logarithmic Curve:

a = Constant =

$$\bar{y} - \frac{b}{n}\sum_{i=1}^{n} \ln x\,(y_i - y)$$

b = lnx coefficient =

$$\frac{\sum_{i=1}^{n} U_{\ln x}(y_i - \bar{y})}{\sum_{i=1}^{n} (U_{\ln x})^2}$$

r^2 = Square of the correlation coefficient =

$$\frac{\sum_{i=1}^{n} U_{\ln x}\,U_x}{\sum_{i=1}^{n} U_{\ln x} U_y}$$

Program 3: One-Way Analysis Of Variance

Group = Analysis of Variance classification.
k = Number of groups.
$\mathbf{n_j}$ = Number of observations in group j, j=1, 2, . . ., k.
$\mathbf{x_{ij}}$ = i^{th} observation in group j, i=1, 2, . . ., n_j; j=1 2, . . ., k.
$\mathbf{x_j}$ = Mean of the j^{th}group, j=1, 2, . . ., k.

x = Overall mean = $\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij} \sum_{j=1}^{k} n_j$

SST = Total corrected Sum of Squares = $\sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2$

SSB = Sum of Squares due to variation Between groups =

$$\sum_{j=1}^{k} n_j \bar{x}_j^2 - (\bar{\bar{x}}^2)(\sum_{j=1}^{k} n_j)$$

ν_1 = Degrees of freedom for between mean square = k−1

MSB = Mean Square due to variation Between groups =

$$\frac{SSB}{\nu_1}$$

SSW = Sum of Squares due to variation Within groups =

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}^2 - \sum_{j=1}^{k} n_j \bar{x}_j^2$$

ν_2 = Degrees of freedom for within mean square =

$$\sum_{j=1}^{k} (n_j - 1)$$

MSW = Mean Square due to variations Within groups =

$$\frac{SSW}{\nu_2}$$

F = Calculated F ratio with ν_1 degrees of freedom for the numerator and ν_2 degrees of freedom for the denominator =

$$\frac{MSB}{MSW}$$

Program 4: Normal Curve Overlay

n = Number of observations

$\mathbf{x_1, x_2 \ldots, x_n}$ = Observations 1, 2, . . . , n

w = Cell width for histogram

x = Sample mean =

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

s = Sample standard deviation =

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

$\mathbf{y_i}$ = Height of normal curve at z_i =

$$\frac{nw}{s\sqrt{2\pi}} e - \frac{w^2}{2s^2} (z_i - z)^2$$

Appendix C
References

Applied Regression Analysis, N. R. Draper and H. Smith, John Wiley & Sons, Inc., New York, 1966.

Statistical Methods, G. W. Snedecor and W. G. Cochran, Iowa State University Press, Ames, Iowa, 1967.

Elements of Statistical Interence, D. V. Huntsberger, Allyn and Bacon, Boston, Mass., 1967.

Biometry, R. R. Sokol and F. J. Rohlf, W. H. Freeman and Co., San Francisco, 1969.

Operating Notes

N O T E 00	Internal Program encountered an error - see instructions on restarting the program
N O T E 04	Too many (keys pressed - parentheses can be nested 5 levels deep
N O T E 05	Too many) keys pressed
N O T E 09	Attempt to find the $\log_{10}$ or Ln value of a negative number. Negative data value entered for program 2.
9.999999999 99R N O T E 11	RANGE OF CALCULATION EXCEEDED - Also check the current result of the accumulator
N O T E 12	Attempt to find $\log_{10}$ or Ln of '0'. '0' entered as data for Program 2.
N O T E 15	Attempt to raise '0' to a negative power
N O T E 16	Attempt to divide by '0'
N O T E 20	Internal Program encountered an error - see instructions on restarting the program
N O T E 21	Internal Program encountered an error - check peripheral device and interconnecting cables
N O T E 22	) pressed before a number entry - enter number and continue
N O T E 24	Not enough) keys pressed.

PART NO. 09805-90004
MICROFICHE NO. 09805-99004

PRINTED IN U.S.A.
APRIL 1973