HEWLETT-PACKARD

HP-85

GENERAL STATISTICS PAC

HEWLETT PACKARD

HP-85

General Statistics Pac

November 1979

00085-90053

Printed in U.S.A.

Introduction

The five sections of the General Statistics Pac have been drawn from the fields of general statistics and related areas. Each section has been designed to fulfill different statistical needs.

Each program in this pac is described by a section in this manual. The manual provides a description of the program with relevant equations, a set of user instructions for using the program, and one or more example problems. Program listings can be obtained by loading the desired program and then listing it. The appendix at the back of this manual provides instructions for obtaining the explanatory comments to a particular program which are stored in the comments program.

If you have already worked through a few programs in the Standard Pac, you will understand how to load a program and how to interpret the User Instructions. If these procedures are not clear to you, take a few minutes to review the sections, "Loading a Program" and "Format of User Instructions", in your Standard Pac.

As stated in the Standard Pac, you should define the output peripherals to your needs. Most of the programs assume that the printer is 2 and the CRT is 1 and use `PRINT` and `DISP` statements accordingly. If you want to ensure that the peripherals are defined as the programs assume, press RESET before running a program. The currently defined key labels are obtainable at any time while a program is running by pressing KEY LABEL. Remember to press CLEAR LINE before pressing END LINE if the key labels are in the input line.

We hope that the General Statistics Pac will assist you in the solution of numerous problems in your discipline.

Contents

Program Operation Hints

These programs have been designed to execute with a minimum amount of difficulty, but problems may occur which you can easily solve during program operation. There are four different types of errors or warnings that can occur while executing a program; input errors, math errors, tape errors and image format string errors.

The input errors include errors 43, 44, and 45. These errors will cause a message to be output followed by a new question mark as a prompt for the input. You should verify your mistake and then enter the correct input. The program will not proceed until the input is acceptable. There is a complete discussion of INPUT in your Owner's Manual if you need more detail.

The second type of error which might occur is a math error (1 thru 13). With DEFAULT ON, the first eight errors listed in Appendix E of your Owner's Manual cause a warning message to be output, but program execution will not be halted. The cause of these errors can usually be attributed to specific characteristics of your data and the type of calculations being performed. In most cases, there is no cause for alarm, but you should direct your attention to a possible problem. An example of such a case is found in the Standard Pac when the curve fitting program computes a curve fit to your data which has a value of 1 for the coefficient of determination, r^2. The computation of the F ratio results in a divide by zero, Warning 8.

The third type of error, tape errors (60 thru 75) may be due to several different problems. Some of the most likely causes are the tape being write-protected, the wrong cartridge (or no cartridge) being inserted, a bad tape cartridge, or wrong data file name specification during program execution. Appendix E of your Owner's Manual should be consulted for a complete listing.

The fourth type of error is due to generalizing the output to anticipated data ranges. In many cases, the output has assumed ranges which may or may not be appropriate with your data. Adjusting the image format string for your data will solve this type of problem. You may also want to change the image string if you require more digits to the right of the decimal point.

These are the more common problems which may occur during program operation. Your Owner's Manual should be consulted if you need more assitance.

One Sample Analysis

This program calculates basic statistics and order statistics for up to 200 data points. The program will also plot a histogram on the CRT and performs a chi-square goodness-of-fit test for three distributions, (normal, exponential, and uniform). In addition, the program performs the t-test and calculates the right-tail area under the t distribution.

Program flow is controlled by use of the special function keys to provide you with the ability to obtain the output you want in the order you want. Once data has been entered from either the keyboard or a tape file, you can print it, edit it, or store it on tape. After the data is entered and edited, you can perform any of the analysis routines by pressing the desired special function key.

Equations, formulae and explanations:

Given a set of data points $\{x_1, x_2, \ldots, x_n\}$

the program computes the following statistics:

mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

standard deviation: $S_x = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n * \bar{x}^2}{(n-1)}}$

standard error of the mean: $S_{\bar{x}} = S_x/\sqrt{n}$

coefficient of variation: $V = S_x/\bar{x} * 100\%$

variance: $S_x{}^2 = \frac{\sum_{i=1}^{n} x_i{}^2 - n * \bar{x}^2}{(n-1)}$

1st moment: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

2nd moment: $m_2 = \frac{1}{n}\Sigma x_i^2 - \bar{x}^2$

3rd moment: $m_3 = \frac{1}{n}\Sigma x_i^3 - \frac{3}{n} x \Sigma x_i^2 + 2\bar{x}^3$

4th moment: $m_4 = \frac{1}{n}\Sigma x_i^4 - \frac{4}{n} x \Sigma x_i^3 + \frac{6}{n} x^2 \Sigma x_i^2 - 3\bar{x}^4$

Moment coefficient of skewness: $\gamma^2 = \dfrac{m_3}{m_2^{3/2}}$

Moment coefficient of kurtosis: $\gamma^2 = \dfrac{m_4}{m_2^2}$

Range $= u_n - u_1$ where u_n is the largest x_i and u_1 is the smallest x_i.

Median x: After the data is ordered, the median x is given by the $n + 1/2^{nd}$ value. Therefore, when n is odd the median is the middle value and when x is even it is the mean of the two middle values.

.25 Quantile: Assuming ordered data, the .25 quantile is given by the $n + 1/4^{th}$ value which may be interpolated from two adjacent values.

.75 Quantile: Assuming ordered data, the .75 quantile is given by the $3(n+1)/4^{th}$ value which may be interpolated from two adjacent values.

Midspread = .75 Quantile − .25 Quantile

Trimean = $x/2$ + (.25 Quantile + .75 Quantile)/4

t Test is a test of the null hypothesis that a random sample comes from a normal population with the mean $\bar{x} = \bar{x}_0$. $t = \sqrt{n}\,(\bar{x} - \bar{x}_0)/S$.

Chi-Square Goodness-of-Fit Test is a test of the null hypothesis that a probability function provides the appropriate model for the data.

$$\chi^2 = \sum_{i=1}^{K} \frac{(f_i - e_i)^2}{e_i}$$

where $f_1, f_2, \ldots, f_k$ are the class frequencies of an observed distribution, while $e_1, e_2, \ldots,$ and e_k are the corresponding frequencies expected for a given distribution, i.e., normal, exponential, or uniform. The validity of the generated statistic is dependent upon the selection of an offset which does not ignore cells which could have non-negligible probabilities.

User Instructions

1. Insert the General Statistics Pac Cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "ONESAM"`.
   b. Press: (END LINE).
3. To start the program:
   a. Press: (RUN).
4. When the keys are labelled and `SELECT OPTION` is displayed:

REFERENCES:

1. *Dictionary/Outline of Basic Statistics,* John D. Freund and Frank J. Williams, McGraw-Hill Book Co., 1966.
2. Abramowitz, M. and Stegun, I.A., *Handbook of Mathematical Functions,* National Bureau of Standards, 1970.

a. Press KEY #5 (HELP), if you need a more detailed explanation.
b. After the explanation is displayed, go to step 4.
OR:
a. Press: KEY #1 (ENTER) to enter the data.

5. When PRINT DATA ON INPUT: Y/N? is displayed:
   a. Enter: Y, if the data is to be printed on entry.
   b. Press: (END LINE)
   OR:
   a. Enter: N, if the data is not to be printed.
   b. Press: (END LINE)
6. When ENTER FROM KEYBOARD/ TAPE: K/T? is displayed:
   a. Enter: K, if the data is to be entered from the keyboard.
   b. Press: (END LINE)
   c. Go to step 7.
   OR:
   a. Enter: T, if the data is on a tape file.
   b. Press: (END LINE)
   c. Go to step 11.
7. When ENTER NUMBER OF POINTS? is displayed:
   a. Enter: The number of points.
   b. Press: (END LINE)

   Note: If the number of points is not in the range of 2 to 200, go to step 7 and enter a number in the acceptable range.
8. When X___)=? is displayed:
   a. Enter: The value specified.
   b. Press: (END LINE)
9. Repeat step 8 for each value.
10. Go to step 12.
11. When ENTER FILE NAME? is displayed:
   a. Enter: The file name.
   b. Press: (END LINE)
   c. After the data has been loaded, go to step 12.

   Note: If the file does not exist, go to step 11.
12. When the data has been entered and DONE is displayed, select the desired option using the specified function keys:
   a. Press: KEY #2 (OUTPUT) to output the array values to the printer or to the tape.
   b. Go to step 13.
   OR:
   a. Press: KEY #3 (PLOT) to plot the histogram.
   b. Go to step 39.
   OR:
   a. Press: KEY #4 (BASIC) to have the basic statistics printed.
   b. Go to step 19.
   OR:
   a. Press: KEY #6 (EDIT) to edit the data in memory.
   b. Go to step 22.
   OR:
   a. Press: KEY #7 (t TEST) to perform a one sample t test or compute a t probability.
   b. Go to step 28.
   OR:
   a. Press: KEY #8 (CHI-SQR) to perform a chi-square goodness-of-fit test.
   b. Go to step 33.
13. When PRINT DATA: Y/N? is displayed:
   a. Enter: Y, if the data is to be printed.
   b. Press: (END LINE)
   c. Go to step 14.
   OR:
   a. Enter: N, if the data is not to be printed.
   b. Press: (END LINE)
   c. Go to step 15.

   Note: You must enter either "Y" or "N" or the program will beep and go to step 13.

14. When PRINT DATA ON PRINTER/DISP: P/D? is displayed:
    a. Enter: P, if the data is to be printed on the printer.
    b. Press: END LINE
    OR:
    a. Enter: D, if the data is to be printed on the display.
    b. Press: END LINE
15. When STORE DATA: Y/N? is displayed:
    a. Enter: Y, to store the data.
    b. Press: END LINE
    c. Go to step 16.
    Note: If NOTHING TO STORE is displayed, go to step 4.
    OR:
    a. Enter: N, to not store the data.
    b. Press: END LINE
    c. Go to step 12.
    Note: You must enter either "Y" or "N" or the program will beep and go to step 15.
16. When ENTER NAME OF FILE? is displayed:
    a. Enter: The file name, (6 characters maximum).
    b. Press: END LINE
17. When CREATE FILE: Y/N? is displayed:
    a. Enter: Y, if the file must be created.
    b. Press: END LINE
    OR:
    a. Enter: N, if the file already exists.
    b. Press: END LINE
    Note: You must enter either "Y" or "N" or the program will beep and go to step 17.
    Note: If any errors occur when storing the data, the program will go to step 16.
18. After the data has been stored, go to step 12.
19. After the basic statistics have been printed and CONF. COEF. FOR C.I. ON μ=? is displayed:
    a. Enter: The confidence coefficient between .7 and .995.
    b. Press: END LINE
    OR:
    a. Enter: 0, if the confidence interval is not desired.
    b. Press: END LINE
    Note: If OUT OF BOUNDS is displayed, go to step 19 and enter a valid coefficient.
    Note: The data is assumed to be normally distributed.
20. When ORDER STATS: Y/N? is displayed:
    a. Enter: Y, if order statistics are desired.
    b. Press: END LINE
    OR:
    a. Enter: N, if the order statistics are not desired.
    b. Press: END LINE
21. When RANKS: Y/N? is displayed:
    a. Enter: Y, if the ranked data are desired.
    b. Press: END LINE
    c. Go to step 12.
    OR:
    a. Enter: N, if the ranked data are not desired.
    b. Press: END LINE
    c. Go to step 12.
    Note: After printing either the order statistics or the ranked data, the data has been sorted into ascending order and can be printed by pressing **KEY #2** (OUTPUT) and requesting a printout.
22. When 0=OK, 1=CORRECT, 2=DELETE, 3=INSERT? is displayed:
    a. Enter: 0, if the edit is finished.
    b. Press: END LINE
    c. Go to step 12.
    OR:
    a. Enter: 1, if you want to correct a data item.
    b. Press: END LINE
    c. Go to step 23.

OR:

a. Enter: 2, if you want to delete a data item.

b. Press: (END LINE)

c. Go to step 25.

OR:

a. Enter: 3, if you want to insert a data item.

b. Press: (END LINE)

c. Go to step 26.

23. When ENTER INDEX OF ITEM TO CORRECT? is displayed:

a. Enter: The index of the item.

b. Press: (END LINE)

c. Go to step 24.

Note: If the index is greater than the number in the data set, go to step 23 or re-enter the index.

OR:

a. Enter: A value less than 1 to terminate the correction mode.

b. Press: (END LINE)

c. Go to step 22.

24. When NEW X (___)=? is displayed:

a. Enter: The correct value.

b. Press: (END LINE)

c. Go to step 22.

25. When ENTER INDEX OF ITEM TO DELETE? is displayed:

a. Enter: The index of the item.

b. Press: (END LINE)

c. Go to step 22.

Note: If the index is greater than the number of items in the data set, go to step 25 and re-enter the index.

OR:

a. Enter: A value less than 1 to terminate the deletion mode.

b. Press: (END LINE)

c. Go to step 22.

26. When ENTER INDEX OF ITEM TO INSERT? is displayed:

a. Enter: The index of the item.

b. Press: (END LINE)

c. Go to step 27.

Note: If the index is greater than the number of items in the data set plus one, go to step 26 and re-enter the index.

Note: If MAXIMUM NUMBER OF ITEMS=200 is displayed, go to step 22 since there is no more room.

OR:

a. Enter: A value less than 1 to terminate the insertion mode.

b. Press: (END LINE)

c. Go to step 22.

27. When INSERT X (___)=? is displayed:

a. Enter: The X value.

b. Press: (END LINE)

c. Go to step 22.

28. When t-TEST OR t-PROB:T/P? is displayed:

a. Enter: T, if a test is desired.

b. Press: (END LINE)

OR:

a. Enter: P, if the probability of a t value with specified number of degrees of freedom is desired.

b. Press: (END LINE)

c. Go to step 31.

Note: You must enter either a "T" or "P" or the program will beep and go to step 28.

29. When 1 OR 2 TAIL TEST? is displayed:

a. Enter: The test type 1 or 2 tail.

b. Press: (END LINE)

Note: You must enter either 1 or 2 or the program will beep and go to step 29.

30. When H0:μ=___OR=? is displayed:

a. Enter: The value for μ.

b. Press: (END LINE)

c. After the results are printed, go to step 12.

31. When `t=?` is displayed:
    a. Enter: The t-value.
    b. Press: END LINE
32. When `DF=?` is displayed:
    a. Enter: The number of degrees of freedom.
    b. Press: END LINE
    c. After the probability is printed, go to step 12.
33. When the goodness-of-fit codes are printed and `GOF CODE ?` is displayed:
    a. Enter: 1, if a normal distribution goodness-of-fit test is desired.
    b. Press: END LINE
    c. Go to step 35.
    OR:
    a. Enter: 2, if an exponential distribution goodness-of-fit test is desired.
    b. Press: END LINE
    c. Go to step 35.
    Note: If `SAMPLE MEAN =___ MUST BE >0 FOR CODE 2` is displayed, the program returns to step 33.
    OR:
    a. Enter: 3, if a uniform distribution goodness-of-fit test is desired.
    b. Press: END LINE
    c. Go to step 34.
    Note: If the entered code is not equal to 1, 2 or 3, the program returns to step 33.
34. When `LOWER & UPPER LIM. ON UNIF.=?` is displayed:
    a. Enter: The lower and upper limits separated by a comma.
    b. Press: END LINE

    Note: If the maximum sample value is lower than the entered lower limit, the program returns to step 34.
35. When `OFFSET=?` is displayed:
    a. Enter: The offset.
    b. Press: END LINE

    Note: If `OFFSET TOO BIG;MAX VALUE=___` is displayed, the program returns to step 35.

    Note: The offset # of cells, and cell width must be chosen so that the expected frequency outside of these bounds is negligible.
36. When `# OF CELLS ?` is displayed:
    a. Enter: The number of cells.
    b. Press: END LINE

    Note: If `# OF CELLS OUT OF BOUNDS:(1,50)` is displayed, the program returns to step 36.
37. When the optimum cell width is displayed and `CELL WIDTH ?` appears in the display:
    a. Enter: The desired cell width.
    b. Press: END LINE
38. If `___OBS. TOO SMALL FOR OFFSET` and/or `___OBS. TOO LARGE FOR ___ CELLS` is displayed and `OFFSET & CELL WIDTH OK:Y/N?` appears in the display:
    a. Enter: Y, if the data is OK.
    b. Press: END LINE
    c. After the test is done, go to step 12.
    OR:
    a. Enter: N, if the data is not OK.
    b. Press: END LINE
    c. Go to step 35.
39. After the program has chained to `ONESM2`, the keys are relabelled, and `OFFSET=?` is displayed:
    a. Enter: The desired offset.
    b. Press: END LINE

    Note: If `OFFSET TOO BIG: MAX VALUE=___` is displayed, go to step 39 and enter a smaller offset.
40. When `# OF CELLS?` is displayed:
    a. Enter: The number of cells.
    b. Press: END LINE

Note: If # OF CELLS OUT OF BOUNDS:(1,50) is displayed, go to step 40 and enter a valid number of cells.

41. When OPTIMUM CELL WIDTH=___ CELL WIDTH? is displayed:
  a. Enter: The desired cell width.
  b. Press: (END LINE)

Note: The optimum cell width is given by:

$$\frac{X_{max} - \text{offset}}{\#\text{ of cells}}$$

42. If ___OBS. TOO SMALL FOR OFFSET and/or ___OBS. TOO LARGE FOR OFFSET is displayed and then OFFSET & CELL WIDTH OK:Y/N? is displayed:
  a. Enter: Y, if the specified information is acceptable.
  b. Press: (END LINE)
  OR:
  a. Enter: N, if you want to change the offset or cell width.
  b. Press: (END LINE)
  c. Go to step 39.

Note: This case results when there are observations, i.e., data values, which fall outside of the specified range. (Offset, Offset + # Cells * Cell Width). The histogram will still be drawn if you want it.

43. When the histogram has been drawn and the program beeps, select the desired option using the specified function keys:
  a. Press: KEY #1 (ENTRY) to chain to the data entry and statistics program ONESAM.
  b. Go to step 4.
  OR:
  a. Press: KEY #3 (PLOT) to plot the histogram.
  b. Go to step 39.
  OR:
  a. Press: KEY #4 (NORMAL) to have a normal curve overlaid over the plot.
  b. Go to step 43.
  OR:
  a. Press: KEY #7 (LABEL) to label the plot at an entered position.
  b. Go to step 44.
  OR:
  a. Press: KEY #8 (COPY) to print the cell statistics and copy the histogram to the printer.
  b. Go to step 43.

44. When LABEL ORIGIN:X,Y? is displayed:
  a. Enter: The X and Y coordinates where the label is to start.
  b. Press: (END LINE)

Note: If INVALID POSITION is displayed, the entered coordinates are out of the scale limits and the program goes to step 44.

Note: To aid label positioning the following variables may be useful to use.

Variable Name	Description
X0	Minimum X-scaled value
Y0	Minimum Y-scaled value
O	Offset
N9	Maximum number of points in a cell
C	Cell Width
D1	Distance of a dot in X-direction
D2	Distance of a dot in Y-direction

45. When ENTER LABEL? is displayed:
  a. Enter: The label.
  b. Press: (END LINE)

Note: If LABEL TOO LONG is displayed, the program beeps and goes to step 45.

46. After the label has been drawn, go to step 43.

Example:

A packager wants a statistical analysis of the weights of 25 boxes of raisins taken from the production of a filling machine to check that the machine is operating within desired limits. The weights of the 25 boxes are:

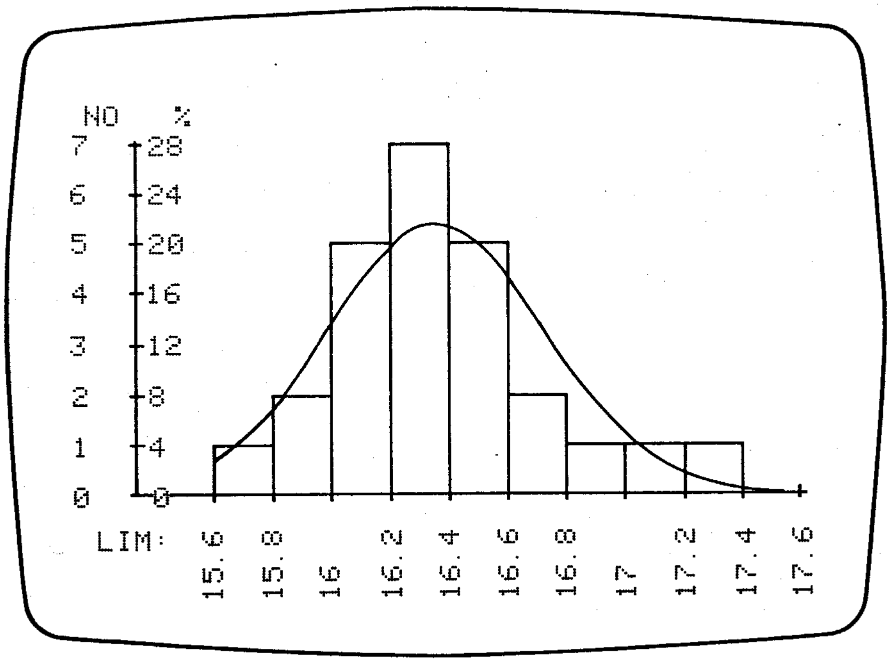
Weight	Weight
16.07	16.38
16.22	16.25
15.68	16.11
17.25	16.41
17.11	16.25
16.59	15.87
16.18	16.41
16.30	16.10
16.02	16.20
15.95	16.61
16.53	16.48
16.89	16.74
16.37	

Enter the data and generate the basic statistics. Use .95 as the confidence coefficient. Perform a 1 tail t test with $\mu = 16.2$. After the statistics are printed, plot a histogram with an offset of 15.60, 10 cells, and a cell width of .2.

A normal curve should be overlaid over the data followed by copying the cell statistics.

```
 I        X(I)          X(I+1)
  1      16.0700        16.2200
  3      15.6800        17.2500
  5      17.1100        16.5900
  7      16.1800        16.3000
  9      16.0200        15.9500
 11      16.5300        16.8900
 13      16.3700        16:3800
 15      16.2500        16.1100
 17      16.4100        16.2500
 19      15.8700        16.4100
 21      16.1000        16.2000
 23      16.6100        16.4800
 25      16.7400
BASIC STATISTICS
********************************
N =    25
STD ERROR OF THE MEAN=          .07
MEAN =          16.3588
COEF OF VARIATION =         2.25%
VARIANCE =           .1354
STANDARD DEVIATION =         .3680
SKEWNESS =           .6516
KURTOSIS =          3.2807
```

```
    95.00% C.I. FOR MEAN:
(     16.2069,      16.5107)
ONE-TAIL t( 24 , .025 )=
 2.06438540174

ORDER STATISTICS
******************************
N =  25
XMIN=   15.6800  XMAX=   17.2500

RANGE=      1.57 MEDIAN=     16.30
.25 QUANTILE=      16.1100
.75 QUANTILE=      16.5300
MID-SPREAD=         .4200
TRIMEAN=      16.3100

RANKED DATA
******************************
RANK    DATA PT. RANK    DATA PT.
   1      15.6800   2      15.8700
   3      15.9500   4      16.0200
   5      16.0700   6      16.1000
   7      16.1100   8      16.1800
   9      16.2000  10      16.2200
  12      16.2500  13      16.3000
  14      16.3700  15      16.3800
  17      16.4100  18      16.4800
  19      16.5300  20      16.5900
  21      16.6100  22      16.7400
  23      16.8900  24      17.1100
  25      17.2500

 1  TAIL t-TEST
HO:μ= 16.2
N=  25
MEAN=      16.3588
STD ERROR OF MEAN=        .0736
t=        2.1577 DF=     24
  P(t>  2.1577)=      .0206

OFFSET=                  15.6
# OF CELLS=              10
CELL WIDTH= .2
```



CELL_STATISTICS

CELL#	LOWER LIMIT	NUMBER OF OBS.	%RELATIVE FREQUENCY
1	15.60	1	4.00
2	15.80	2	8.00
3	16.00	5	20.00
4	16.20	7	28.00
5	16.40	5	20.00
6	16.60	2	8.00
7	16.80	1	4.00
8	17.00	1	4.00
9	17.20	1	4.00

Paired Sample Analysis

This program calculates basic statistics on paired data, performs polynomial and family regression, and plots the X-Y pairs and the regression curves. The program also calculates the paired t statistic and the right-tail area under the t distribution. Four nonparametric routines are also included. Two measures of the ranks of the samples are available: Spearman's Rho and Kendall's Tau. Two tests are available to measure the difference between the two groups: the sign test and Wilcoxon signed rank test. Both tests eliminate the effect of pairing in order to measure the true group difference.

Program flow is controlled by use of the special function keys to provide you with the ability to obtain the output you want in the order you want. Once data has been entered from either the keyboard or a tape file, you can print it, edit it, or store it on tape. After the data has been entered and edited, you can perform any of the analysis routines by pressing the desired special function key.

Equations, formulae, and explanations:

means: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

standard deviations: $s_x = \sqrt{\frac{\Sigma x_i^2 - n\bar{x}^2}{n-1}}$ Variance $x = s_x^2$

$s_y = \sqrt{\frac{\Sigma y_i^2 - n\bar{y}^2}{n-1}}$ Variance $y = s_y^2$

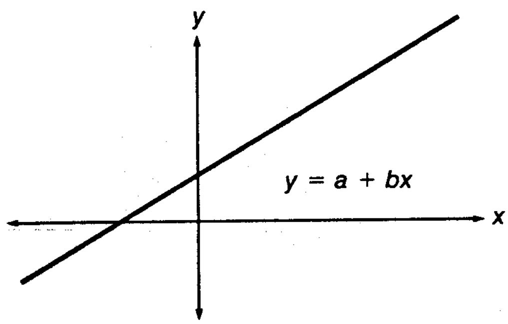
correlation coefficient: $r_{xy} = \dfrac{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)}{s_x * s_y}$

polynomial curve fit: $y = b_0 + b_1 x + b_2 x^2 + \ldots + b_k x^k$

For a polynomial function fit by the method of least squares the values of $b_0, b_1, b_2, \ldots, b_k$ are obtained by solving the system of $k + 1$ normal equations.

$$\Sigma y = n\, b_0 + b_1 (\Sigma x) + b_2 (\Sigma x^2) + \ldots + b_k (\Sigma x^k)$$
$$\Sigma xy = b_0 (\Sigma x) + b_1 (\Sigma x^2) + b_2 (\Sigma x^3) + \ldots + b_k (\Sigma x^{k+1})$$
$$\vdots \qquad \vdots$$
$$\Sigma x^k y = b_0 (\Sigma x^k) + b_1 (\Sigma x^{k+1}) + b_2 (\Sigma x^{k+2}) + \ldots + b_k (\Sigma x^{2k})$$

Linear Regression

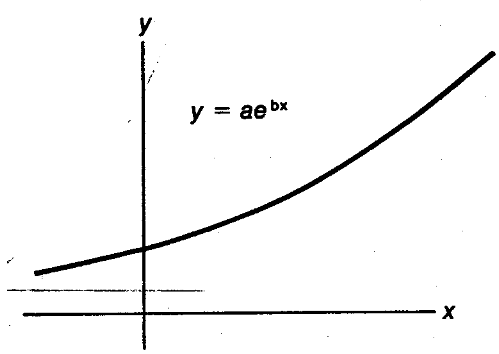


$$b = \frac{\Sigma x_i y_i - \dfrac{\Sigma x_i \, \Sigma y_i}{n}}{\Sigma x_i^2 - \dfrac{(\Sigma x_i)^2}{n}}$$

$$a = \left[\frac{\Sigma y_i}{n} - b\,\frac{\Sigma x_i}{n}\right]$$

$$r^2 = \frac{\left[\Sigma x_i y_i - \dfrac{\Sigma x_i \, \Sigma y_i}{n}\right]^2}{\left[\Sigma x_i^2 - \dfrac{(\Sigma x_i)^2}{n}\right]\left[\Sigma y_i^2 - \dfrac{(\Sigma y_i)^2}{n}\right]}$$

Exponential Curve Fit

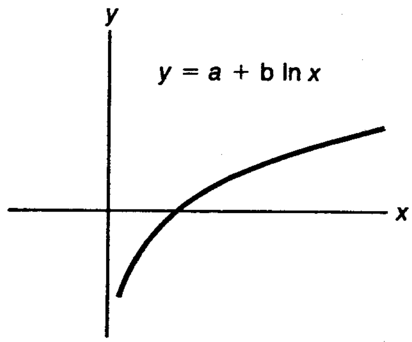


$$b = \frac{\Sigma x_i \ln y_i - \dfrac{1}{n}(\Sigma x_i)(\Sigma \ln y_i)}{\Sigma x_i^2 - \dfrac{1}{n}(\Sigma x_i)^2}$$

$$a = \exp\left[\frac{\Sigma \ln y_i}{n} - b\,\frac{\Sigma x_i}{n}\right]$$

$$r^2 = \frac{\left[\Sigma x_i \ln y_i - \dfrac{1}{n}\Sigma x_i \, \Sigma \ln y_i\right]^2}{\left[\Sigma x_i^2 - \dfrac{(\Sigma x_i)^2}{n}\right]\left[\Sigma (\ln y_i)^2 - \dfrac{(\Sigma \ln y_i)^2}{n}\right]}$$

Logarithmic Curve Fit

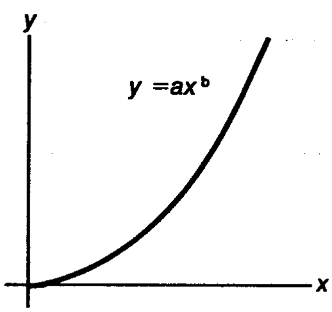


$$b = \frac{\Sigma y_i \ln x_i - \frac{1}{n}\Sigma \ln x_i \; \Sigma y_i}{\Sigma(\ln x_i)^2 - \frac{1}{n}(\Sigma \ln x_i)^2}$$

$$a = \frac{1}{n}(\Sigma y_i - b\,\Sigma \ln x_i)$$

$$r^2 = \frac{\left[\Sigma y_i \ln x_i - \frac{1}{n}\Sigma \ln x_i \; \Sigma y_i\right]^2}{\left[\Sigma(\ln x_i)^2 - \frac{1}{n}(\Sigma \ln x_i)^2\right]\left[\Sigma y_i^2 - \frac{1}{n}(\Sigma y_i)^2\right]}$$

Power Curve Fit



$$b = \frac{\Sigma(\ln x_i)(\ln y_i) - \frac{(\Sigma \ln x_i)(\Sigma \ln y_i)}{n}}{\Sigma(\ln x_i)^2 - \frac{(\Sigma \ln x_i)^2}{n}}$$

$$a = \exp\left[\frac{\Sigma \ln y_i}{n} - b\,\frac{\Sigma \ln x_i}{n}\right]$$

$$r^2 = \frac{\left[\Sigma(\ln x_i)(\ln y_i) - \frac{(\Sigma \ln x_i)(\Sigma \ln y_i)}{n}\right]^2}{\left[\Sigma(\ln x_i)^2 - \frac{(\Sigma \ln x_i)^2}{n}\right]\left[\Sigma(\ln y_i)^2 - \frac{(\Sigma \ln y_i)^2}{n}\right]}$$

Paired t test is an application of the one-sample t test of differences between paired data which are assumed to be normal. The assumption of equal standard deviations in the two-sample t test cannot be made.

let

$\bar{D}_i = x_i - y_i$ which is normal

and

$\Delta = \mu_1 - \mu_2$

then

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n}\bar{D}_i$$

$$S_{\bar{D}} = \frac{\bar{D}_i^2 - \frac{1}{n}(\Sigma\bar{D}_i)^2}{n-1}$$

$$S_{\bar{D}} = S_{\bar{D}}/\sqrt{n}$$

the test statistic: $t = \frac{\bar{D}}{S_{\bar{D}}}$ which has $n-1$ degrees of freedom (df) can be used to test the null hypothesis H_0: $\mu_1 - \mu_2 = \Delta$ (difference)

Nonparametrics: This section assumes that the x_i data is paired or related to the y_i data. In order to calculate the correlation between the X's and Y's, two measures of correlation on the ranks of the original data are available. Spearman's Rho determines the ranks for the X's and the ranks for the Y's and then calculates the ordinary correlation coefficient on the ranks. The Kendall's Tau rank correlation is slightly more complicated.

Spearman's Rho

From the N pairs, the measure of rank correlation given by ρ, where

$$d = \sum_{i=1}^{N} \left(R_{X_i} - R_{Y_i} \right)^2 \qquad \text{where } \rho = 1 - \frac{6d}{N(N^2 - 1)}$$

R_{X_i} is the rank of X_i among the X's and R_{Y_i} is the rank of Y_i among the Y's.

Kendall's Tau

From the N pairs is determined the number of concordant pairs, P_c, and the number of discordant pairs, P_d, for the permutations of N pairs taken two pairs at a time.

A pair $((X_i, Y_i)$ and $(X_j, Y_j))$ is concordant if $X_i - X_j > 0$ and $Y_i - Y_j > 0$ or $X_i - X_j < 0$ and $Y_i - Y_j < 0$. If $X_i = X_j$ or $Y_i = Y_j$, the pair is disregarded, and the pair is discordant otherwise. The test statistic is given by

$$\tau = \frac{P_i - P_d}{N(N - 1)/2}$$

Two tests are available to determine if the X group (say treatment one) is significantly different from the Y group (treatment two). Both tests eliminate the effects of the pairing of the observations in order to measure true group difference. The two tests and assorted hypotheses are:

The Sign Test

$$\text{H}_0\text{: Prob.}\left[X > Y\right] = \text{P}\left[X < Y\right] = \frac{1}{2}$$

Essentially, this test says that if the X and Y samples come from the same population, and if we look at the differences between paired observations, then one-half of the signs of these differences should be positive and one-half negative. This test really is a binomial test equivalent to determining if a coin is balanced (one-half heads, one-half tails). From the number of pairs, N, is obtained the number of positive differences $X_i - Y_i$, T, and a normalized T based on T being distributed binomally under the hypothesis. Points where $X_i = Y_i$ are excluded from the analysis reducing the total N.

The normalized T is given by

$$\frac{T - \frac{N}{2}}{\sqrt{\frac{N}{4}}}$$

Wilcoxon Signed Rank Test

H_0: The values of X tend to be about the same as the values of Y. The average values are about the same.

This test is an equivalent test to the paired t test. Instead of using only the signs of the differences, as in the sign test, this test ranks the differences and forms the sum of the signed ranks. It is, in most applications, a more powerful test than the sign test, i.e., a better test.

The N differences, $X_i - Y_i$, are ranked from smallest to largest. T, the test statistic, is given by the sum of the ranks of the positive differences. Pairs for which $X_i = Y_i$ are excluded from the analysis.

Two standard normal deviates based on the conditional distribution of the existing ties are output 1) with the correction for continuity, $\frac{1}{2}$, and 2) without the correction for continuity. The standard normal deviates are derived from:

$$\frac{T = \mu_T}{\sigma_T} \text{ where}$$

$$\mu_T = \frac{N(N+1)}{4}, \quad \text{or} \quad \mu_T = \frac{N(N+1)}{4} - \frac{1}{2}$$

and $\sigma_T^2 = \frac{1}{4} \sum_{j=1}^{k} \frac{1}{t_j} \sum_{I=S_{j-1}+1}^{S_j} I$, where I is the rank.

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: LOAD "PAIRED"
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and SELECT OPTION is displayed:
   a. Press: KEY #5 (HELP), if you need a more detailed explanation.
   b. After the explanation is displayed, go to step 4.

   OR:

   a. Press: KEY #1 (ENTER), to enter the data pairs.
   b. Go to step 5.

   OR:

   a. Go to step 14, if data has been entered.

REFERENCES:

1. Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Functions*, National Bureau of Standards, 1970.
2. Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering*, John Wiley & Sons, 1965.
3. Conover, W. J., *Practical Nonparametric Statistics*, John Wiley & Sons, 1971.
4. Graybill, F. A., *An Introduction to Linear Statistical Models, Vol. I*, McGraw-Hill, 1961.
5. Mosteller, F. and Rourke, Robert E. K., *Sturdy Statistics*, Addison-Wesley, 1973.
6. Siegel, S., *Nonparametric Statistics*, McGraw-Hill, 1956.

5. When `PRINT DATA ON INPUT: Y/N?` is displayed:
   a. Enter: Y, if you want the data printed on entry.
   b. Press: END LINE
   OR:
   a. Enter: N, if you do not want the data printed on entry.
   b. Press: END LINE

   Note: You must enter either Y or N or the program will beep and go to step 5.
6. When `ENTER FROM KEYBOARD/ TAPE: K/T?` is displayed:
   a. Enter: K, if the data is to be entered from the keyboard.
   b. Press: END LINE
   OR:
   a. Enter: T, if the data is to be entered from an existing data file.
   b. Press: END LINE
   c. Go to step 12.

   Note: You must enter either K or T or the program will beep and go to step 6.
7. When `NO. OF POINTS?` is displayed:
   a. Enter: The number of data pairs.
   b. Press: END LINE

   Note: The maximum number of pairs is 200.
8. If `INVALID NUMBER OF POINTS` is displayed, go to step 7 and re-enter a valid number.
9. When `X(__),Y(__)=?` is displayed:
   a. Enter: The X and Y values separated by a comma.
   b. Press: END LINE
10. Repeat step 9 for each pair.
11. When `DATA ENTERED` is displayed, go to step 13.
12. When `ENTER FILE NAME?` is displayed:
    a. Enter: The file name.
    b. Press: END LINE

    Note: If an error occurs during the load, the program returns to step 12.
13. When `MAX DEGREE REGRESSION?` is displayed:
    a. Enter: The maximum regression degree for polynomial regression.
    b. Press: END LINE

    Note: If the entered degree is greater than 9 and `1<=MAX DEG<=9` is displayed, the program returns to step 13.

    Note: Higher maximum regression degrees may lead to matrix instability when computing the preliminary analysis of variance for the polynomial regression.
14. After `DONE` is displayed, any of the options can be selected.
    a. Press: KEY #2 (`OUTPUT`), to output the data to either the printer or a tape file.
    b. Go to step 15.
    OR:
    a. Press: KEY #3 (`BASIC`), to have the basic statistics printed.
    b. After the basic statistics are printed, go to step 14.
    OR:
    a. Press: KEY #4 (`REGRESS`), to perform any of the regressions or plot the X - Y data.
    b. Go to step 36.
    OR:
    a. Press: KEY #6 (`EDIT`), to edit the data.
    b. Go to step 20.
    OR
    a. Press: KEY #7 (`NONPARA`), to obtain any of the nonparametric statistics.
    b. Go to step 62.
    OR:
    a. Press: KEY #8 (`t TEST`), to perform a paired t test or compute a t probability.
    b. Go to step 28.

15. When PRINT DATA:Y/N? is displayed:
    a. Enter: Y, to print the data.
    b. Press: END LINE

    OR

    a. Enter: N, if a printout of the data is not wanted.
    b. Press: END LINE

    Note: You must enter either Y or N or the program will beep and go to step 15.
16. When STORE DATA:Y/N? is displayed:
    a. Enter: Y, if you want to store the data.
    b. Press: END LINE
    c. Go to step 17.

    OR:

    a. Enter: N, if you do not want to store the data.
    b. Press: END LINE
    c. Go to step 19.

    Note: You must enter either Y or N or the program will beep and go to step 16.
17. When ENTER NAME OF FILE? is displayed:
    a. Enter: The file name.
    b. Press: END LINE
18. When CREATE FILE:Y/N? is displayed:
    a. Enter: Y, to create the file.
    b. Press: END LINE

    OR:

    a. Enter: N, if the file already exists.
    b. Press: END LINE

    Note: You must enter either Y or N or the program will beep and go to step 18.
19. When DONE is displayed, go to step 14.
20. When 0=OK,1=CORRECT,2=DELETE,3=INSERT? is displayed:
    a. Enter: 0, if the edit is finished.
    b. Press: END LINE
    c. Go to step 27.

    OR:

    a. Enter: 1, if you want to correct a data pair.
    b. Press: END LINE
    c. Go to step 21.

    OR:

    a. Enter: 2, if you want to delete a data pair.
    b. Press: END LINE
    c. Go to step 24.

    OR:

    a. Enter: 3, if you want to insert a data pair.
    b. Press: END LINE
    c. Go to step 25.
21. When INDEX OF PAIR TO CORRECT? is displayed:
    a. Enter: The index of the data pair.
    b. Press: END LINE
    c. Go to step 22.

    Note: If the index is greater than the number of data pairs in the data set, go to step 21 and re-enter the index.

    OR:

    a. Enter: A value less than 1 to terminate the correction mode.
    b. Press: END LINE
    c. Go to step 20.
22. When NEW X(__)=? is displayed:
    a. Enter: The correct value.
    b. Press: END LINE
23. When NEW Y(__)=? is displayed:
    a. Enter: The correct value.
    b. Press: END LINE
    c. Go to step 20.
24. When ENTER INDEX OF PAIR TO DELETE? is displayed:
    a. Enter: The index of the pair.
    b. Press: END LINE
    c. Go to step 20.

    Note: If the index is greater than the number of data pairs in the data set, go to step 24 and re-enter the index.

    OR:

    a. Enter: A value less than 1 to terminate the deletion mode.

b. Press: END LINE
c. Go to step 20.

25. When ENTER INDEX OF PAIR TO INSERT? is displayed:
a. Enter: The index of the pair.
b. Press: END LINE
c. Go to step 26.

Note: If the index is greater than the number of data pairs in the data set plus one, go to step 25 and re-enter the index.

Note: If MAXIMUM NUMBER OF PAIRS=200 is displayed, go to step 20 since there is no more room.

OR:
a. Enter: A value less than 1 to terminate the insertion mode.
b. Press: END LINE
c. Go to step 20.

26. When INSERT X(__), Y(__)=? is displayed:
a. Enter: The X and Y values separated by a comma.
b. Press: END LINE
c. Go to step 20.

27. When DONE is displayed, go to step 13.

28. When t-TEST OR t-PROB:T/P? is displayed:
a. Enter: T, if a t test is desired.
b. Press: END LINE

OR:
a. Enter: P, if the t probability of a t value with a specified number of degrees of freedom is desired.
b. Press: END LINE
c. Go to step 34.

Note: You must enter T or P or the program will beep and go to step 28.

29. When 1 OR 2 TAILED? is displayed:
a. Enter: 1, for a 1 tailed test.
b. Press: END LINE
c. Go to step 30.

OR:
a. Enter: 2, for a 2 tailed test.
b. Press: END LINE
c. Go to step 31.

Note: You must enter "1" or "2" or the program will beep and go to step 29.

30. When H0:μ(X) - μ(Y)=? is displayed:
a. Enter: The difference.
b. Press: END LINE
c. Go to step 32.

31. When H0:ABS(μ(X) - μ(Y))=? is displayed:
a. Enter: The difference.
b. Press: END LINE

32. When LEVEL OF SIGNIFICANCE? is displayed:
a. Enter: The level of significance which is in the range of (.005, .3).
b. Press: END LINE

Note: If the level of significance is not in bounds and OUT OF BOUNDS (.005, .3) is displayed, the program returns to step 32.

33. When t PROB:Y/N? is displayed:
a. Enter: Y, if the t probability is desired for the t value.
b. Press: END LINE
c. When DONE is displayed, go to step 14.

OR:
a. Enter: N, if the t probability is not desired.
b. Press: END LINE
c. When DONE is displayed, go to step 14.

Note: You must enter Y or N or the program will beep and go to step 33.

34. When t? is displayed:
a. Enter: The t value.
b. Press: END LINE

35. When DF? is displayed:

a. Enter: The degrees of freedom.
b. Press: END LINE
c. When DONE is displayed, go to step 14.

36. After the program has chained to the regression program the keys are labelled and PLOT DATA:Y/N? is displayed:
a. Enter: Y, if the data pairs are to be plotted.
b. Press: END LINE
c. Go to step 37.
OR:
a. Enter: N, if the data is not to be plotted.
b. Press: END LINE
c. Go to step 48.

37. When AUTO X-SCALING:Y/N? is displayed:
a. Enter: Y, if the X-minimum and X-maximum values are to be used by the program.
b. Press: END LINE
c. Go to step 40.
OR:
a. Enter: N, if you want to enter the X-minimum and X-maximum values.
b. Press: END LINE

Note: By specifying the end points you can have better control of the axis labels.

Note: You must enter either Y or N or the program will beep and go to step 37.

38. When ENTER SCALE XMIN? is displayed:
a. Enter: The minimum X value for scaling.
b. Press: END LINE

39. When ENTER SCALE XMAX? is displayed:
a. Enter: The maximum X value for scaling.
b. Press: END LINE

Note: If the maximum value is less than or equal to the minimum value, the program will beep and go to step 38.

40. When VERTICAL/HORIZONTAL LABELS:V/H? is displayed:
a. Enter: V, if the X-axis labels are to be written vertically.
b. Press: END LINE
OR:
a. Enter: H, if the X-axis labels are to be written horizontally.
b. Press: END LINE

Note: You must enter V or H or the program will beep and go to step 40.

41. When NO. OF X-AXIS INTERVALS:(<=16)? is displayed:
a. Enter: The number of X-axis intervals (≤16).
b. Press: END LINE

Note: If the number of intervals is not in the range of 1 to 16, the program will beep and go to step 41.

42. When NUMBER OF X-INT. BETWEEN LABELS? is displayed:
a. Enter: The number of X-intervals between labels, e.g., if labels are desired at every other tic, the number of intervals between labels is 2.
b. Press: END LINE
OR:
a. Enter: 0, if no labels are desired on the X-axis.
b. Press: END LINE

Note: If the number of intervals is not in the range of 0 to the entered number of X-intervals, the program will beep and go to step 42.

43. When AUTO Y-SCALING:Y/N? is displayed:
a. Enter: Y, if the Y-minimum and Y-maximum values are to be used by the program.
b. Press: END LINE
c. Go to step 46.
OR:
a. Enter: N, if you want to enter the Y-minimum and Y-maximum values.

b. Press: END LINE

Note: By specifying the end points you can have better control of the axis labels.

Note: You must enter either Y or N or the program will beep and go to step 43.

44. When ENTER SCALE YMIN? is displayed:
   a. Enter: The minimum Y value for scaling.
   b. Press: END LINE

45. When ENTER SCALE YMAX? is displayed:
   a. Enter: The maximum Y value for scaling.
   b. Press: END LINE

   Note: If the maximum value is less than or equal to the minimum value, the program will beep and go to step 44.

46. When NO. OF Y-AXIS INTERVALS: (<=12)? is displayed:
   a. Enter: The number of Y-axis intervals (≤12).
   b. Press: END LINE

   Note: If the number of intervals is not in the range of 1 to 12, the program will beep and go to step 46.

47. When NUMBER OF Y-INT. BETWEEN LABELS? is displayed:
   a. Enter: The number of Y-intervals between labels, e.g., if labels are desired at every other tic, the number of intervals between labels is 2.
   b. Press: END LINE

   OR:

   a. Enter: 0, if no labels are desired on the Y-axis.
   b. Press: END LINE

   Note: If the number of intervals is not in the range of 0 to the entered number of Y-intervals, the program will beep and go to step 47.

48. After SELECT OPTION is displayed, any of the options can be selected.

a. Press: KEY #1 (CHAIN) to chain to data entry or nonparametrics.
b. Go to step 49.

OR:

a. Press: KEY #2 (POLY), to obtain a polynomial curve fit.
b. Go to step 50.

Note: If MATRIX UNSTABLE-USE SMALLER MAXIMUM DEGREE! is displayed: go to step 48. The current matrix is unstable and you cannot perform a polynomial regression with the currently defined maximum degree value. To perform a polynomial regression of any degree now, you must chain to the enter program, edit the data base, and then enter a maximum degree regression at step 13 which is lower than the current value.

OR:

a. Press: KEY #6 (XY-PLOT), to obtain an XY-PLOT of the data and specify the plotting option.
b. Go to step 37.

OR:

a. Press: KEY #3 (LINEAR), to obtain the linear curve fit.
b. Go to step 51.

OR:

a. Press: KEY #4 (EXP), to obtain the exponential curve fit.
b. Go to step 51.

OR:

a. Press: KEY #7 (LOG), to obtain the logarithmic curve fit.
b. Go to step 51.

OR:

a. Press: KEY #8 (POWER), to obtain the power curve fit.
b. Go to step 51.

Note: If CAN'T TAKE LOG is displayed, the data contains values less than or equal to 0 and this regression cannot be done. The program returns to step 48.

49. When CHAIN ENTRY/ NONPARAMETRICS:E/N? is displayed:
    a. Enter: E, if you want to chain to the data entry program.
    b. Press: (END LINE)
    c. Go to step 4.
    OR:
    a. Enter: N, if you want to chain to the nonparametrics program.
    b. Press: (END LINE)
    c. Go to step 62.
    Note: You must either enter E or N or the program will beep and go to step 49.
50. When DEGREE REGRESSION? is displayed:
    a. Enter: The degree of the desired regression curve.
    b. Press: (END LINE)
    Note: If MAX REG = ___ is displayed, go to step 50 and enter a degree less than or equal to the maximum degree entered at step 13.
51. When ESTIMATE Y:Y/N? is displayed
    a. Enter Y, if ŷ is to be calculated.
    b. Press: (END LINE)
    c. Go to step 52.
    OR:
    a. Enter: N, if no estimates are desired.
    b. Press: (END LINE)
    c. Go to step 56.
    Note: You must enter either Y or N or the program will beep and go to step 51.
52. When AT ALL X(I):Y/N? is displayed:
    a. Enter: Y, if estimates of Y at all entered X values are desired.
    b. Press: (END LINE)
    c. Go to step 53.
    OR:
    a. Enter: N, if all are not desired.
    b. Press: (END LINE)
    c. Go to step 54.
    Note: You must enter either Y or N or the program will beep and go to step 52.
53. When ESTIMATE Y AT ENTERED X:Y/N? is displayed:
    a. Enter: Y, to obtain an estimate of Y.
    b. Press: (END LINE)
    c. Go to step 54.
    OR:
    a. Enter: N, if no more estimates are desired.
    b. Press: (END LINE)
    c. Go to step 56.
    Note: You must enter either Y or N or the program will beep and go to step 53.
54. When ESTIMATE Y AT X=? is displayed:
    a. Enter: The X value for computing the Y estimate.
    b. Press: (END LINE)
55. When ANOTHER ESTIMATE:Y/N? is displayed:
    a. Enter: Y, if another estimate is desired.
    b. Press: (END LINE)
    c. Go to step 54.
    OR:
    a. Enter: N, if no more estimates are desired.
    b. Press: (END LINE)
    c. Go to step 56.
    Note: You must enter either Y or N or the program will beep and go to step 55.
56. If plotting was selected during entry and PLOT:Y/N? is displayed:
    a. Enter: Y, if the regression curve is to be plotted.
    b. Press: (END LINE)

c. Go to step 58.
OR:
a. Enter: N, if the plot is not wanted.
b. Press: END LINE
c. Go to step 48.

Note: You must enter either Y or N or the program will beep and go to step 56.

57. If plotting was not done during entry, go to step 48.
58. When LABEL PLOT:Y/N? is displayed:
a. Enter: Y, if you want to label the plot.
b. Press: END LINE
c. Go to step 59.
OR:
a. Enter: N, if no label is desired.
b. Press: END LINE
c. Go to step 48.

Note: You must enter either Y or N or the program will beep and go to step 58.

59. When LABEL ORIGIN:X,Y? is displayed:
a. Enter: The X and Y coordinates where the label is to start.
b. Press: END LINE

Note: If INVALID POSITION is displayed, the entered coordinates are out of the scale limits and the program goes to step 59.

Note: To aid label positioning the following variables may be useful to use.

Variable Name	Description
X0	Minimum X-scaled value
Y0	Minimum Y-scaled value
X1	X-value at axes intercept
Y1	Y-value at axes intercept
X2	X-value at right end of X-axis
Y2	Y-value at top end of Y-axis
Z1	X2—X1
Z2	Y2—Y1
D1	Distance of a dot in X-direction
D2	Distance of a dot in Y-direction

60. When ENTER LABEL? is displayed:
a. Enter: The label.
b. Press: END LINE

Note: If LABEL TOO LONG is displayed, the program beeps and goes to step 60.

61. After the label has been drawn, go to step 58.
62. After the program has chained to PAIR3 and and SELECT OPTION is displayed:
a. Press: KEY #1 (ENTER), to chain to the data entry program.
b. Go to step 4.
OR:
a. Press: KEY #2 (SPEARMAN), to calculate Spearman's Rho.
b. After Spearman's Rho is printed, go to step 62.
OR:
a. Press: KEY #3 (SIGN), to perform a sign test.
b. After the results are printed, go to step 62.
OR:
a. Press: KEY #4 (REGRESS), to perform any of the regressions or plot the X-Y data.
b. Go to step 36.
OR:
a. Press: KEY #5 (HELP), if you need a more detailed explanation.
b. After the explanation is displayed, go to step 62.
OR:
a. Press: KEY #6 (KENDALL), to calculate Kendall's Tau.
b. After Kendall's Tau is printed, go to step 62.
OR:
a. Press: KEY #7 (WILCOXON), to perform a Wilcoxon signed rank test.
b. After the results are printed, go to step 62.

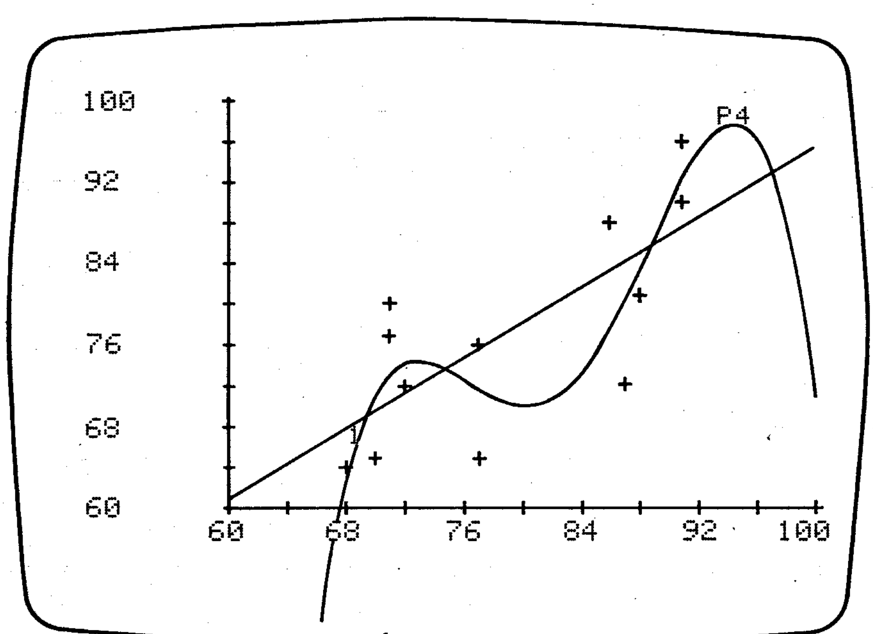
Example:

The following data has been selected to demonstrate the program and is not from either two paired samples or two independent samples. Using the data listed below, obtain basic statistics, a polynomial regression of degree 4, a linear regression, and the four nonparametric test results.

I	1	2	3	4	5	6	7	8	9	10	11	12
X(I)	86	71	77	68	91	72	77	91	70	71	88	87
Y(I)	88	77	76	64	96	72	65	90	65	80	81	72

```
  I         X(I)         Y(I)
  1       86.0000      88.0000
  2       71.0000      77.0000
  3       77.0000      76.0000
  4       68.0000      64.0000
  5       91.0000      96.0000
  6       72.0000      72.0000
  7       77.0000      65.0000
  8       91.0000      90.0000
  9       70.0000      65.0000
 10       71.0000      80.0000
 11       88.0000      81.0000
 12       87.0000      72.0000

MAXIMUM DEGREE REGRESSION=   4

BASIC STATISTICS
*******************************
MEANS,VARIANCES,CORRELATION
X MEAN=      79.0833
VAR(X)=      78.9924
Y MEAN=      77.1667
VAR(Y)=     107.6061
   MINIMUM        MAXIMUM
    68.0000         91.0000
    64.0000         96.0000
RXY =          .7344


   PRELIMINARY AOV TABLE

SOURCE     SS      F       DF    R^2
TOTAL   72640.000
MEAN    71456.333
TOT ADJ 1183.667
X^1   638.47   11.71 (1,  10) .539
X^2    53.06     .97 (1,   9) .584
X^3    95.01    1.91 (1,   8) .664
X^4    31.87     .61 (1,   7) .691
```

```
DEGREE REGRESSION =  4

   ANALYSIS OF VARIANCE
SOURCE/DF      SS       MS      F
TOTAL 11    1183.7
REG    4     818.4   204.6    3.9
RESID  7     365.3    52.2
R SQUARE =          0.691


COEFFICIENTS
 I      B(I)     STD ERROR     T

 0 -43211.48  49630.37        -.87
 1   2135.79   2509.26         .85
 2    -39.31     47.41        -.83
 3       .32       .40         .81
 4      -.00       .00        -.78

   AOV: LINEAR REG:CODE 1
SOURCE/DF      SS       MS      F
TOTAL 11    1183.7
REG    1     638.5   638.5  11.7
RESID 10     545.2    54.5
R SQUARE =          0.539


YHAT =      9.377+       0.857 X
```



```
SPEARMAN'S RHO
---------- ---
   I  RANK OF X(I)  RANK OF Y(I)

     1       8.0000      10.0000
     2       3.5000       7.0000
     3       6.5000       6.0000
     4       1.0000       1.0000
     5      11.5000      12.0000
     6       5.0000       4.5000
     7       6.5000       2.5000
     8      11.5000      11.0000
     9       2.0000       2.5000
    10       3.5000       8.0000
    11      10.0000       9.0000
    12       9.0000       4.5000
SUM OF SQUARED RANK DIFFERENCES=
    75.0000
RHO=        .7378


KENDALL'S TAU
--------- ---

NUMBER OF CONCORDANT PAIRS=  49
NUMBER OF DISCORDANT PAIRS=  12

TAU=       .5606


SIGN TEST
---- ----

N=    11
NO. OF POSITIVE DIFFERENCES=  7
(THE  1 POINTS WHERE X(I) = Y(I)
 ARE EXCLUDED FROM THE TEST)

YIELDS APPROX. STD. NOR. DEV.=
       .9045


WILCOXON SIGNED RANK
-------- ------ ----


N=   11
SUM OF POSITIVE RANKS=  41.5

(USING RANKS OF X(I)-Y(I) AND
EXCLUDING THE  1
POINTS WHERE X(I)=Y(I))
```

```
YIELDS APPROXIMATE STANDARD
NORMAL DEVIATES
 1) WITHOUT CORRECTION FOR
    CONTINUITY:

A)NOT COMPENSATING FOR

TIED DIFFERENCES:
                .7557

B)CONDITIONAL ON THE

EXISTING TIED DIFFERENCES:
                .7565
 2) WITH CORRECTION FOR
    CONTINUITY

A)NOT COMPENSATING FOR

TIED DIFFERENCES:
                .7113

B)CONDITIONAL ON THE

EXISTING TIED DIFFERENCES:
                .7120
```

Test Statistics

This section containing five programs performs a chi-square test for equal and unequal expected values, calculates chi-square probability, performs a t test for means of two samples, calculates chi-square for r × c (row by column) contingency tables, and performs one-way and two-way analysis of variance. Each of the five programs employs the special function keys to control program flow. The following table shows the five programs and the test statistics included in each one.

Program Name	Test Statistic
CHISQR	Chi-square test for equal and unequal expected values
	Chi-square probability calculation
RXCCHI	Chi-square for r × c (row by column) contingency tables
	Chi-square probability calculation
TSTAT2	t test for the means of two samples
	t probability calculation
ONEAOV	One-way analysis of variance
	Chi-square probability calculation
TWOAOV	Two-way analysis of variance with one or more replications.

Each program contains routines for data entry, output, and editing in addition to the statistics.

REFERENCES:

1. Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Functions,* National Bureau of Standards, 1970.
2. Dixon and Massey, *Introduction to Statistical Analysis,* McGraw-Hill, 1969.
3. Freund, J. E., *Mathematical Statistics,* Prentice Hall, 1962.
4. Ostle, B., *Statistics in Research,* Iowa State University Press, 1972.

Chi-Square Tests

With unequal expected values,

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

where

O_i = Observed frequency
E_i = Expected frequency
n = Number of cells

If the expected values are equal,

$$\left(E = E_i = \frac{\Sigma O_i}{n} \text{ for all } i \right)$$

then

$$\chi_2^2 = \frac{n\Sigma O_i^2}{\Sigma O_i} - \Sigma O_i$$

Chi-square probability function with N degrees of freedom

$$f(x) = \frac{1}{\Gamma(N/2)^2} N/2 * x^{N/2-1} e^{-x/2}$$

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "CHISQR"`
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and `SELECT OPTION` is displayed:
   a. Press: KEY #5 (`HELP`), if you need a more detailed explanation.
   b. After the explanation is displayed, go to step 4.

   OR:

   a. Press: KEY #1 (`CHI=`), to enter data for a chi-square test with equal expected values.
   b. Go to step 5.

OR:

a. Press KEY #2 (CHI#), to enter data for a chi-square test with unequal expected values.

b. Go to step 10.

OR:

a. Press: KEY #4 (CHI² PROB), to enter a chi-square value and number of degrees of freedom and obtain the probability.

b. Go to step 30.

5. When MAX SIZE = 300: # OF CELLS, K=? is displayed:

   a. Enter: The number of cells.

   b. Press: (END LINE)

   Note: If the number of cells is not in the range of 1 to 300, the program will beep and go to step 5.

6. When PRINT DATA ON ENTRY: Y/N? is displayed:

   a. Enter: Y, if the data is to be printed.

   b. Press: (END LINE)

   OR:

   a. Enter: N, if the data is not to be printed.

   b. Press: (END LINE)

   Note: You must enter either Y or N or the program will beep and go to step 6.

7. When OBS. FREQ. FOR CELL ___? is displayed:

   a. Enter: The frequency for the specified cell.

   b. Press: (END LINE)

8. Repeat step 7 for each cell.

9. Go to step 15.

10. When MAX SIZE = 300: # OF CELLS, K=? is displayed:

    a. Enter: The number of cells.

    b. Press: (END LINE)

    Note: If the number of cells is not in the range of 1 to 300, the program will beep and go to step 10.

11. When PRINT DATA ON ENTRY: Y/N? is displayed:

    a. Enter: Y, if the data is to be printed on entry.

    b. Press: (END LINE)

    OR:

    a. Enter: N, if the data is not to be printed.

    b. Press: (END LINE)

    Note: You must either enter Y or N or the program will beep and go to step 11.

12. When OBS. FREQ. FOR CELL ___? is displayed:

    a. Enter: The observed frequency for the specified cell.

    b. Press: (END LINE)

13. When EXP. FREQ. FOR CELL ___? is displayed:

    a. Enter: The expected frequency for the specified cell.

    b. Press: (END LINE)

14. Repeat steps 12 and 13 for each cell.

15. When DONE is displayed:

    a. Press: KEY #3 (TEST) to perform the chi-square test on the entered data.

    b. After the test is done, go to step 15.

    Note: If YOU MUST ENTER DATA FIRST! is displayed, go to step 4 and enter data before trying this option.

    OR:

    a. Press: KEY #6 (EDIT) to edit the data in memory.

    b. Go to step 16.

    Note: If YOU MUST ENTER DATA FIRST! is displayed, go to step 4 and enter data before trying this option.

    OR:

    a. Press: KEY #7 (PRINT) to print the data currently in memory.

    b. After the data is printed, go to step 15.

    Note: If YOU MUST ENTER DATA FIRST! is displayed, go to step 4 and enter data before trying this option.

OR:

a. Go to step 4 and enter a new set of data.

16. When `EDITING CODE?` is displayed:
 a. Enter: 0, if the edit is finished.
 b. Press: (END LINE)
 c. Go to step 15.

 OR:

 a. Enter: 1, if you want to add a cell.
 b. Press: (END LINE)
 c. Go to step 17.

 Note: If `CANNOT ADD ANY MORE` is displayed, go to step 16 since there is no more room.

 OR:

 a. Enter: 2, if you want to delete a cell.
 b. Press: (END LINE)
 c. Go to step 21.

 OR:

 a. Enter: 3, if you want to change the value(s) of a cell.
 b. Press: (END LINE)
 c. Go to step 24.

 Note: If you enter a value out of the range of 0 to 3, the program will beep and go to step 16.

17. If the data has unequal expected values, go to step 19.

18. When `OBS. FREQ. FOR CELL__?` is displayed:
 a. Enter: The frequency for the specified cell.
 b. Press: (END LINE)
 c. Go to step 16.

19. When `OBS. FREQ. FOR CELL__?` is displayed:
 a. Enter: The observed frequency for the specified cell.
 b. Press: (END LINE)

20. When `EXP. FREQ. FOR CELL__?` is displayed:
 a. Enter: The expected frequency for the specified cell.
 b. Press: (END LINE)
 c. Go to step 16.

21. If the data has unequal expected values, go to step 23.

22. When `DELETE O(I) WHERE I=?` is displayed:
 a. Enter: The index of the cell to delete.
 b. Press: (END LINE)
 c. Go to step 16.

 OR:

 a. Enter: 0, to terminate the deletion mode.
 b. Press: (END LINE)
 c. Go to step 16.

 Note: If the index is out of the range of 0 to the number of items in the data set, go to step 22 and re-enter the index.

23. When `DELETE O(I),E(I) WHERE I=?` is displayed:
 a. Enter: The index of the cell to delete.
 b. Press: (END LINE)
 c. Go to step 16.

 OR:

 a. Enter: 0, to terminate the deletion mode.
 b. Press: (END LINE)
 c. Go to step 16.

 Note: If the index is out of the range of 0 to the number of items in the data set, go to step 23 and re-enter the index.

24. If the data has unequal expected values, go to step 27.

25. When `CHANGE O(I) WHERE I=?` is displayed:
 a. Enter: The index of the cell to change.
 b. Press: (END LINE)
 c. Go to step 26.

 OR:

 a. Enter: 0, to terminate the change mode.
 b. Press: (END LINE)
 c. Go to step 16.

Note: If the index is out of the range of 0 to the number of items in the data set, go to step 25 and re-enter the index.

26. When `NEW O(__)=?` is displayed:
    a. Enter: The frequency for the specified cell.
    b. Press: (END LINE)
    c. Go to step 16.

27. When `CHANGE O(I),E(I) WHERE I=?` is displayed:
    a. Enter: The index of the cell to change.
    b. Press: (END LINE)
    c. Go to step 28.

    OR:

    a. Enter: 0, to terminate the change mode.
    b. Press: (END LINE)
    c. Go to step 16.

Note: If the index is out of the range of 0 to the number of items in the data set, go to step 27 and re-enter the index.

28. When `NEW O(__)=?` is displayed:
    a. Enter: The observed frequency for the specified cell.
    b. Press: (END LINE)

29. When `NEW E(__)=?` is displayed:
    a. Enter: The expected frequency for the specified cell.
    b. Press: (END LINE)
    c. Go to step 16.

30. When `DF=?` is displayed:
    a. Enter: The degrees of freedom.
    b. Press: (END LINE)

31. When `CHI-SQUARE=?` is displayed:
    a. Enter: the χ^2 value.
    b. Press: (END LINE)
    c. After the probability is printed, go to step 4.

χ^2 Test Examples:

1. Perform a χ^2 test with unequal expected values using the following data:

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

```
CHI-SQUARE '#' EXPECTED VALUES
     I        O(I)        E(I)
     1         8.0000       9.6000
     2        50.0000      46.7500
     3        47.0000      51.8500
     4        56.0000      54.4000
     5         5.0000       8.2500
     6        14.0000       9.1500
K= 6

   I    OBSERVED         EXPECTED
        FREQUENCY        FREQUENCY
    1          8.00           9.60
    2         50.00          46.75
    3         47.00          51.85
    4         56.00          54.40
    5          5.00           8.25
    6         14.00           9.15

CHI-SQUARE=      4.8444
K= 6
DF= 5
PROB CHI-SQUARE >      4.8444
                 =      .4352
```

2. The following table shows the observed frequencies in tossing a die 120 times. Perform a χ^2 test using this data with equal expected value.

number	1	2	3	4	5	6
frequency o_i	25	17	15	23	24	16

```
CHI-SQUARE '=' EXPECTED VALUES

       I       O(I)
        1      25.0000
        2      17.0000
        3      15.0000
        4      23.0000
        5      24.0000
        6      16.0000
K= 6

   I    OBSERVED         EXPECTED
        FREQUENCY        FREQUENCY
    1        25.00           20.00
    2        17.00           20.00
    3        15.00           20.00
    4        23.00           20.00
    5        24.00           20.00
    6        16.00           20.00

CHI-SQUARE=     5.0000
K= 6
DF= 5
PROB CHI-SQUARE >      5.0000
                   =    .4159
```

3. What is the χ^2 probability if χ^2 value is 4.2 and the number of degrees of freedom is 5.

```
CHI-SQUARE PROBABILITY
DF= 5
CHI-SQUARE= 4.2
PROB CHI-SQUARE >      4.2000
                   =    .5210
```

RxC Contingency Table

i \ j	1	2	3	4	...	C	TOTALS
1	x_{11}	x_{12}	x_{13}	x_{14}		x_{1C}	R_1
2	x_{21}						
3	x_{31}						
.	.						
.	.						
.	.						
R	x_{R1}					x_{RC}	R_R
TOTALS	C_1	C_2				C_c	T

This program computes the χ^2 statistic with $(R-1)(C-1)$ (degrees of freedom) for testing the independence of the two variables. Also Pearson's coefficient of contingency C_c, which measures the degree of association between the two variables, is calculated.

Row sum $$R_i = \sum_{j=1}^{C} x_{ij} \quad i = 1, 2, \ldots, R$$

Column sum $$C_j = \sum_{i=1}^{R} x_{ij} \quad j = 1, 2, \ldots, C$$

Total $$T = \sum_{i=1}^{R} \sum_{j=1}^{C} x_{ij}$$

Chi-square statistic $$\chi^2 = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{(x_{ij} - E_{ij})^2}{E_{ij}}$$

$$= T \left(\sum_{i=1}^{R} \sum_{j=1}^{C} \frac{x_{ij}^2}{R_i\, C_j} \right) - T$$

$$df = (R - 1)(C - 1)$$

Where the expected frequency $$E_{ij} = \frac{R_i\, C_j}{T}$$

Contingency coefficient $$C_c = \frac{\chi^2}{T + \chi^2}$$

Chi-square probability function with N degrees of freedom

$$f(x) = \frac{1}{\Gamma(N/2)^2} N/2 * x^{N/2-1} e^{-x/2}$$

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "RXCCHI"`
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and `SELECT OPTION` is displayed:
   a. Press: KEY #5 (`HELP`), if you need a more detailed explanation.
   b. After the explanation is displayed, go to step 4.
   OR:
   a. Press: KEY #1 (`ENTER`), to enter the data.
   b. Go to step 5.
   OR:
   a. Press: KEY #4 (`CHI² PROB`), to enter a chi-square value and number of degrees of freedom and obtain the probability.
   b. Go to step 13.
5. When `MAX SIZE = 12: # OF ROWS?` is displayed:
   a. Enter: The number of rows.
   b. Press: (END LINE)

   Note: If the number of rows is not in the range of 2 to 12, the program will beep and go to step 5.
6. When `MAX SIZE = 12: # OF COLUMNS?` is displayed:
   a. Enter: The number of columns.
   b. Press: (END LINE)

   Note: If the number of columns is not in the range of 2 to 12, the program will beep and go to step 6.
7. When `PRINT DATA ON ENTRY: Y/N?` is displayed:
   a. Enter: Y, if the data is to be printed on entry.
   b. Press: (END LINE)
   OR
   a. Enter: N, if the data is not to be printed.
   b. Press: (END LINE)

   Note: You must enter either Y or N or the program will beep and go to step 7.
8. When `CELL (__,__)?` is displayed:
   a. Enter: The value of the specified cell.
   b. Press: (END LINE)
9. Repeat step 8 for each cell.
10. When `DONE` is displayed:
    a. Press: KEY #2 (`PRINT`), to print the data.
    b. After the data is printed, go to step 10.

    Note: If `MUST ENTER DATA FIRST!` is displayed, go to step 4 and enter data before trying this option.

    OR:
    a. Press: KEY #3 (`CHISQR`), to perform a χ^2 test.
    b. After the test is done, go to step 10.

    Note: If `MUST ENTER DATA FIRST!` is displayed, go to step 4 and enter data before trying this option.

    OR:
    a. Press: KEY #6 (`EDIT`), to edit the data.
    b. Go to step 11.

    Note: If `MUST ENTER DATA FIRST!` is displayed, go to step 4 and enter data before trying this option.

    OR:
    a. Go to step 4 and enter another data set.

11. When ROW AND COLUMN OF CELL: R, C? is displayed:
    a. Enter: The row and column indices of the cell to edit.
    b. Press: (END LINE)

    Note: If OUT OF LIMITS: __, __ is displayed, go to step 11 and enter valid indices.

12. When NEW VALUE=? is displayed:
    a. Enter: The cell frequency.
    b. Press: (END LINE)
    c. Go to step 10.

13. When DF=? is displayed:
    a. Enter: The number of degrees of freedom.
    b. Press: (END LINE)

14. When CHI-SQUARE=? is displayed:
    a. Enter: The value for chi-square.
    b. Press: (END LINE)
    c. After the probability is printed, go to step 4.

Example:

A random sample of 250 men and 250 women were polled as to their desires concerning the ownership of television sets. The data in the following table resulted. Apply the program to analyze the result of the poll, i.e., can the hypothesis that the desire to own a television set is independent of sex be rejected?

Results of Sample Poll on Television Ownership

Classification	Men	Women	Total
Want television	80	120	200
Don't want television	170	130	300
Total	250	250	500

```
CHI-SQUARE R X C CONTINGENCY
             TABLE

******INPUT DATA******
ROW  1:     80    120
ROW  2:    170    130


CELL FREQUENCY
ROW  1:     80    120
ROW  2:    170    130
```

```
ROW  1: 80  120
ROW  2: 170  130

TOTALS:
 COLUMN       TOTAL
C( 1) =         250
C( 2) =         250

  ROW         TOTAL
R( 1) =         200
R( 2) =         300


OVERALL=        500

EXPECTED FREQUENCY
ROW   1:
          100.00
          100.00

ROW   2:
          150.00
          150.00


# OF EXP. FREQ. <=2 = 0
% EXP. FREQ. <=5 =   0.00%
CHI-SQUARE =    13.3333
CONTINGENCY COEFFICIENT =
      .1612
DF= 1
PROB CHI-SQUARE >   13.3333
                =    .0003
```

Two Sample t Test

Suppose $\{x_1, x_2, \ldots, x_{n1}\}$ and $\{y_1, y_2, \ldots, y_{n2}\}$ are independent random samples from two normal populations having means μ_1, μ_2 (unknown) and the same unknown variance σ^2.

We want to test the null hypothesis

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

$$x = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i$$

$$y = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i$$

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

We can use this t statistic which has the t distribution with $n_1 + n_2 - 2$ degrees of freedom (df) to test the null hypothesis H_0.

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "TSTAT2"`
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and `SELECT OPTION` is displayed:
   a. Press: KEY #5 (`HELP`), if you need a more detailed explanation.
   b. After the explanation is displayed, go to step 4.
   OR:
   a. Press: KEY #1 (`ENTER`), to enter the data.
   b. Go to step 5.
   OR:
   a. Press: KEY #4 (`t PROB`), to enter a t value and number of degrees of freedom and obtain the probability.
   b. Go to step 17.
5. When `PRINT ON ENTRY:Y/N?` is displayed:
   a. Enter: Y, if the data is to be printed on entry.
   b. Press: (END LINE)
   OR:
   a. Enter: N, if the data is not to be printed.
   b. Press: (END LINE)
   Note: You must enter either Y or N or the program will beep and go to step 5.
6. When `MAX SIZE = 200:# PTS. IN SAMPLE__?` is displayed:
   a. Enter: The number of points in the specified sample.
   b. Press: (END LINE)
   Note: If the number of points is not in the range of 2 to 200, the program beeps and goes to step 6.
7. When `DATA POINT__?` is displayed:
   a. Enter: The value of the specified data point.
   b. Press: (END LINE)
8. Repeat step 7 for each point in the sample.
9. Repeat steps 6, 7 and 8 for each sample.

10. When DONE is displayed:
    a. Press: KEY #2 (PRINT), to print the sample data.
    b. After the data is printed, go to step 10.

    Note: If MUST ENTER DATA is displayed, go to step 4 and enter the data before trying this option.

    OR:
    a. Press: KEY #3 (t TEST), to compute basic statistics and perform a t test on the means of the two samples.
    b. After the t test is done, go to step 10.

    Note: If MUST ENTER DATA is displayed, go to step 4 and enter the data before trying this option.

    OR:
    a. Press: KEY #4 (t PROB), to enter a t value and number of degrees of freedom and obtain the probability.
    b. Go to step 17.

    OR:
    a. Press: KEY #6 (EDIT), to edit the data.
    b. Go to step 11.

    OR:
    a. Go to step 4 to enter a new set of data.
11. When EDIT SAMPLE 1 OR 2? is displayed:
    a. Enter: The sample number.
    b. Press: END LINE

    Note: If you do not enter 1 or 2, the program will beep and go to step 11.
12. When EDIT CODE? is displayed:
    a. Enter: 0, if the edit is finished.
    b. Press: END LINE
    c. Go to step 10.

    OR:
    a. Enter: 1, if you want to add a value to either sample.
    b. Press: END LINE
    c. Go to step 13.

    Note: If MAXIMUM ENTERED is displayed, go to step 12 since there is no more room in this sample.

    OR:
    a. Enter: 2, if you want to delete a value for a sample.
    b. Press: END LINE
    c. Go to step 14.

    OR:
    a. Enter: 3, if you want to change a value in a sample.
    b. Press: END LINE
    c. Go to step 15.

    Note: If you enter a value out of the range of 0 to 3, the program will beep and go to step 12.
13. When DATA POINT ___? is displayed:
    a. Enter: The value of the specified point.
    b. Press: END LINE
    c. Go to step 12.
14. When DELETE X(I) WHERE I=? or DELETE Y(I) WHERE I=? is displayed:
    a. Enter: The index of the point to delete.
    b. Press: END LINE
    c. Go to step 12.

    Note: If the index is not in the range of 1 to the number of points in the sample, the program will beep and go to step 14.
15. When CHANGE X(I) WHERE I=? or CHANGE Y(I) WHERE I=? is displayed:
    a. Enter: The index of the point to change.
    b. Press: END LINE

    Note: If the index is not in the range of 1 to the number of points in the sample, the program will beep and go to step 15.
16. When NEW VALUE=? is displayed:
    a. Enter: The new value.
    b. Press: END LINE
    c. Go to step 12.
17. When t? is displayed:
    a. Enter: the value of t.
    b. Press: END LINE
18. When DF? is displayed:
    a. Enter: The number of degrees of freedom.
    b. Press: END LINE
    c. After the probability is printed, go to step 4.

Example:

Perform a t test using the data from two samples shown below:

Sample 1 $N_1 = 10$	184	22	40	129	26	47	138	42	84	173		
Sample 2 $N_2 = 12$	192	64	235	223	66	224	41	51	152	144	68	186

```
t STATISTIC FOR THE MEANS OF TWO
             SAMPLES


SAMPLE  1
N = .10

      I      X(I)
      1    184.0000
      2     22.0000
      3     40.0000
      4    129.0000
      5     26.0000
      6     47.0000
      7    138.0000
      8     42.0000
      9     84.0000
     10    173.0000


SAMPLE  2
N =  12

      I      Y(I)
      1    192.0000
      2     64.0000
      3    235.0000
      4    223.0000
      5     66.0000
      6    224.0000
      7     41.0000
      8     51.0000
      9    152.0000
     10    144.0000
     11     68.0000
     12    186.0000

N FOR 1 =  10
1 MEAN =  88.5
STD. DEV. FOR 1 =       62.304
```

```
N FOR 2 =   12
2 MEAN =  137.166666667
STD. DEV. FOR 1 =       75.083

t=       1.6325 DF=     20
PROB t >      1.6325 =      .0591
```

One-way Analysis of Variance

The one-way analysis of variance is used to test if observed differences among k sample means can be attributed to chance or whether they are indicative of actual differences among the corresponding population means. Suppose the ith sample has n_i observations (samples may have equal or unequal number of observations). The null hypothesis we want to test is that the k population means are all equal. This program generates the complete ANOVA table.

1. Mean of observations in the i^{th} sample ($i = 1, 2, \ldots, k$)

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

2. Standard deviation of observations in the i^{th} sample

$$s_i = \left[\left(\sum_{j=1}^{n_i} x_{ij}^2 - n_i \bar{x}_i^2 \right) \Big/ (n_i - 1) \right]^{1/2}$$

3. Sum of observations in the i^{th} sample

$$\text{Sum}_i = \sum_{j=1}^{n_i} x_{ij}$$

4. Total sum of squares

$$\text{TSS} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}^2 - \frac{\left(\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij} \right)^2}{\sum_{i=1}^{k} n_i}$$

5. Treatment sum of squares

$$\text{TrSS} = \sum_{i=1}^{k} \frac{\left(\sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i} - \frac{\left(\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij} \right)^2}{\sum_{i=1}^{k} n_i}$$

6. Error sum of squares

$$\mathrm{ESS} = \mathrm{TSS} - \mathrm{TrSS}$$

7. Treatment degrees of freedom

$$\mathrm{df}_1 = k - 1$$

8. Error degrees of freedom

$$\mathrm{df}_2 = \sum_{i=1}^{k} n_i - k$$

9. Total degrees of freedom

$$\mathrm{df}_3 = \mathrm{df}_1 + \mathrm{df}_2 = \sum_{i=1}^{k} n_i - 1$$

10. Treatment mean square

$$\mathrm{TrMS} = \frac{\mathrm{TrSS}}{\mathrm{df}_1}$$

11. Error mean square

$$\mathrm{EMS} = \frac{\mathrm{ESS}}{\mathrm{df}_2}$$

12. The F ratio

$$F = \frac{\mathrm{TrMS}}{\mathrm{EMS}} \text{ (with degrees of freedom } \mathrm{df}_1,\ \mathrm{df}_2)$$

13. Chi-square probability function with N degrees of freedom

$$f(x) = \frac{1}{\Gamma(N/2)^2} N/2 * x^{N/2-1} e^{-x/2}$$

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: LOAD "ONEAOV"
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and SELECT OPTION is displayed:
   a. Press: KEY #5 (HELP), if you need a

Two-way Analysis of Variance

This program analyzes the source of variation in a two way table of data with one or more replications. The number of replications per cell must be equal. Let r be the number of rows, c be the number of columns, and n be the number of replications.

Equations:

1. Sums

Cell $$C_{ij} = \sum_k x_{ijk}$$

Row $$RS_i = \sum_j \sum_k x_{ijk}$$

Column $$CS_j = \sum_i \sum_k x_{ijk}$$

where

$i = 1, 2, \ldots, r$

$j = 1, 2, \ldots, c$

$k = 1, 2, \ldots, n$

2. Sums of squares

Total $$TSS = \sum_i \sum_j \sum_k x_{ijk}^2 - \left(\sum_i \sum_j \sum_k x_{ijk} \right)^2 / nrc$$

Row $$RSS = \sum_i \left(\sum_j \sum_k x_{ijk} \right)^2 / nc - \left(\sum_i \sum_j \sum_k x_{ijk} \right)^2 / nrc$$

Column $$CSS = \sum_j \left(\sum_i \sum_k x_{ijk} \right)^2 / nr - \left(\sum_i \sum_j \sum_k x_{ijk} \right)^2 / nrc$$

Error $$ESS = \sum_i \sum_j \sum_k x_{ijk}^2 - \sum_i \sum_j \left(\sum_k x_{ijk} \right)^2 / n$$

Interaction $$ISS = \sum_i \sum_j \left(\sum_k x_{ijk} \right)^2 / n - \sum_i \left(\sum_j \sum_k x_{ijk} \right)^2 / nc$$

$$- \sum_j \left(\sum_i \sum_k x_{ijk} \right)^2 / nr + \left(\sum_i \sum_j \sum_k x_{ijk} \right)^2 nrc$$

$$= TSS - RSS - CSS - ESS$$

3. Degrees of freedom

$df_1 = r - 1$
$df_2 = c - 1$
$df_3 = (r - 1)(c - 1)$
$df_4 = rc(n - 1)$

4. Mean squares

Row $\quad \text{RMS} = \dfrac{\text{RSS}}{df_1}$

Column $\quad \text{CMS} = \dfrac{\text{CSS}}{df_2}$

Interaction $\quad \text{IMS} = \dfrac{\text{ISS}}{df_3}$

Error $\quad \text{EMS} = \dfrac{\text{ESS}}{df_4}$

5. *F* ratios

Row
- fixed model $\quad F_1 = \dfrac{\text{RMS}}{\text{EMS}}$
- pooled model $\quad F_1' = \dfrac{\text{RMS}}{(\text{IMS} + \text{EMS})}$

Column
- fixed model $\quad F_2 = \dfrac{\text{CMS}}{\text{EMS}}$
- pooled model $\quad F_2' = \dfrac{\text{CMS}}{(\text{IMS} + \text{EMS})}$

Interaction $\quad F_3 = \dfrac{\text{IMS}}{\text{EMS}}$

6. Contrasts

Row $\quad \displaystyle\sum_{i=1}^{r} C_i = 0$

Comparison = $\sum_{i=1}^{r} C_i * \frac{C_{ij}}{(n*c)}$

Contrast SS $CSS_r = \left(\sum_{i=1}^{r} C_i * C_{ij}\right)^2 \Big/ \left(n * c * \left(\sum_{i=1}^{r} C_i^2\right)\right)$

F ratio = $\frac{CSS_r}{EMS}$

Column $\sum_{i=1}^{c} C_i = 0$

Comparison = $\sum_{i=1}^{c} C_i * \frac{C_{ij}}{n*r}$

Contrast SS $CSS_c = \left(\sum_{i=1}^{c} C_i * C_{ij}\right)^2 \Big/ \left(n * r * \left(\sum_{i=1}^{c} C_i^2\right)\right)$

F ratio = $\frac{CSS_c}{EMS}$

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "TWOAOV"`
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and `SELECT OPTION` is displayed:
   a. Press: KEY #5 (`HELP`), if you need a more detailed explanation.
   b. After the explanation is displayed, go to step 4.

   OR:

   a. Press: KEY #1 (`ENTER`), to enter data.
   b. Go to step 5.
5. When `NO. OF ROWS?` is displayed:
   a. Enter: The number of rows.
   b. Press: (END LINE)
6. When `NO. OF COLUMNS?` is displayed:
   a. Enter: The number of columns.
   b. Press: (END LINE)

   Note: If the number of rows times the number of columns is greater than 300, the program will beep and go to step 5.
7. When `NO. OF REPLICATIONS?` is displayed:
   a. Enter: The number of replications.
   b. Press: (END LINE)
8. When `Y(__,__,__)=?` is displayed:
   a. Enter: The specified value.
   b. Press: (END LINE)
9. Repeat step 8 for each replication, column, and row.
10. When `DONE` is displayed:
    a. Press: KEY #2 (`BASIC`), to print basic statistics.
    b. After the statistics are printed, go to step 10.

    Note: If `MUST ENTER DATA` is displayed, you must go to step 4 and enter the data before trying this option.

OR:

a. Press: KEY #3 (AOV), to print an AOV table. If the number of replications is greater than 1, the interaction sum of squares, mean square and F-ratio are printed.

b. After the table is printed, go to step 10.

Note: If MUST ENTER DATA is displayed, you must go to step 4 and enter the data before trying this option.

OR:

a. Press: KEY #4 (POOL AOV), to print an AOV table pooling interaction sum of squares and degrees of freedom, with error sum of squares and degrees of freedom.

b. After the table is printed, go to step 10.

Note: If MUST ENTER DATA is displayed, you must go to step 4 and enter the data before trying this option.

OR:

a. Press: KEY #6 (EDIT), to edit a value.

b. Go to step 11.

Note: If MUST ENTER DATA is displayed, you must go to step 4 and enter the data before trying this option.

OR:

a. Press: KEY #7 (CONTRAST), to print a contrast of either row or column means.

b. Go to step 14.

Note: If MUST ENTER DATA is displayed, you must go to step 4 and enter the data before trying this option.

OR:

a. Press: KEY #8 (TUKEY), to print Tukey's test for interaction if the number of replications is equal to one. If the number of replications is greater than one, this key prints an AOV table as described for KEY #3 (AOV).

b. After the table is printed go to step 10.

Note: If MUST ENTER DATA is displayed, you must go to step 4 and enter the data before trying this option.

11. When CORRECT:ROW#,COL#? is displayed:

a. Enter: The row and column number of the cell to correct.

b. Press: (END LINE)

Note: If R,C MUST BE < =__,__, RESPECTIVELY is displayed, go to step 11 and enter valid values.

12. When ROW __, COL __ WRONG VALUE=? is displayed:

a. Enter: The value of the specified cell which is wrong.

b. Press: (END LINE)

13. When CORRECT VALUE=? is displayed:

a. Enter: The correct value.

b. Press: (END LINE)

c. Go to step 10.

14. When ROW OR COLUMN CONTRASTS:R/C? is displayed:

a. Enter: R, if you want a row contrast printed.

b. Press: (END LINE)

OR:

a. Enter: C, if you want a column contrast printed.

b. Press: (END LINE)

Note: You must enter either R or C or the program will beep and go to step 14.

15. When C( )=? is displayed:

a. Enter: The specified contrast coefficient.

b. Press: (END LINE)

16. Repeat step 15 for each coefficient.

Note: If $\Sigma c_i \neq 0$, then INVALID CONTRAST will be printed.

17. After the results are printed, go to step 10.

Example:

Using the following data set, perform a two way analysis of variance. Do a pooled analysis of variance after the analysis of variance is finished. Print a contrast of row means where $C_1 = 1$, $C_2 = 0$, and $C_3 = -1$.

R= 3		C= 3	N= 2
ROW	COL	OBSERVATIONS	
1	1	3.0000	3.0000
	2	4.0000	3.0000
	3	3.0000	4.0000
2	1	6.0000	4.0000
	2	6.0000	7.0000
	3	8.0000	5.0000
3	1	3.0000	4.0000
	2	6.0000	7.0000
	3	5.0000	6.0000

```
TWO WAY ANALYSIS OF VARIANCE

R=    3
C=    3
N=    2

ROW   COLUMN OBSERVATIONS

   1     1     3.00      3.00
         2     4.00      3.00
         3     3.00      4.00

   2     1     6.00      4.00
         2     6.00      7.00
         3     8.00      5.00

   3     1     3.00      4.00
         2     6.00      7.00
         3     5.00      6.00

ROW   COL   CELL MEAN   CELL VAR
   1     1      3.000      0.000
         2      3.500      0.500
         3      3.500      0.500

   2     1      5.000      2.000
         2      6.500      0.500
         3      6.500      4.500

   3     1      3.500      0.500
         2      6.500      0.500
         3      5.500      0.500
```

```
ROW MEANS:
 ROW      MEAN

   1      3.333
   2      6.000
   3      5.167


COL MEANS:
 COL      MEAN

   1      3.833
   2      5.500
   3      5.167


OVERALL MEAN =     4.833

    ANALYSIS OF VARIANCE

SOURCE/DF      SS      MS      F
TOTAL   17    44.5
ROWS     2    22.3    11.2   10.6
COLS     2     9.3     4.7    4.4
RXC      4     3.3     0.8    0.8
ERROR    9     9.5     1.1


    ANALYSIS OF VARIANCE

SOURCE/DF      SS      MS      F
TOTAL   17    44.5
ROWS     2    22.3    11.2   11.3
COLS     2     9.3     4.7    4.7
ERROR   13    12.8     1.0


   ROW CONTRAST:
   1.0000,   0.0000,  -1.0000,

    COMPARISON =       -1.833
    CONTRAST SS =      10.083
    FRATIO =           10.214
```

Distributions

This program provides algorithms for evaluating exact right-tailed probabilities for the normal, Student's t, central F, chi-square, binomial, Poisson, Weibull, and hypergeometric distributions. All probabilities are calculated with a minimum of five digits of accuracy and, in most cases, seven digits of accuracy. The special function keys are used to select the desired probability.

Wherever possible, continued fraction expansions are used since the convergence tends to be more rapid than for infinite series expansions. The only true approximation used is the normal distribution. All other distributions are approximations only in the sense that the finite continued fractions are approximations to infinite continued fractions.

For the central F, Student's t, and chi-square distributions, the algorithms generally converge most rapidly for small or large right tail probabilities. For moderate tails, the time increases as the right tail approaches .5. For the chi-square, it is recommended that the degrees of freedom be less than 500.

For the normal distribution, a polynomial approximation due to Hastings is used.

Equations, formulae, and explanations:

For the continuous distributions, normal, central F, Student's t, chi-square, and Weibull, the right-tailed probability is defined as $P(x > a)$ where x is a random variable and a is an observable value of x.

For the discrete distributions, binomial, Poisson, and hypergeometric, the right-tailed probability is defined as $P(X \geq x)$ where X is a random variable and x is an element in the counter domain of X.

The distributions are defined as follows:

For the continuous distributions, let f(x) be a density then:

a. Normal $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad -\infty < x < \infty$

b. Central F with N degrees of freedom in the numerator and D in the denominator

$$f(x) = \frac{\Gamma((N+D)/2)\,(N/D)^{N/2}}{\Gamma(N/2)\,\Gamma(D/2)} \quad \frac{x^{N/2-1}}{\left(1 + \frac{N \times x}{D}\right)^{(N+D)/2}}$$

c. Student's t with N degrees of freedom

$$f(x) = \frac{\Gamma((N+1)/2)}{\sqrt{N\pi}\,\Gamma(N/2)} \times \frac{1}{(1 + x^2/N)^{(N+1)/2}}$$
$$-\infty < x < \infty$$

d. Chi-square with N degrees of freedom

$$f(x) = \frac{1}{\Gamma(N/2)\, 2^{N/2}} \times x^{N/2-1} e^{-x/2}$$

e. Weibull with parameters α, β

$$f(x) = -\alpha\beta x^{\beta-1}\exp[-\alpha x^{\beta}]$$
$$x > 0$$

Discrete distributions

f. Binomial

Let N = number of trials
p = probability of success at each trial
X = number of successes

$$P(X = R) = \binom{N}{R} p^R(1-p)^{N-R}$$
$$R = 0, 1, \ldots, N$$

and

$$PO = P(X \geqslant R) = \sum_{i=R}^{N} \binom{N}{i} p^i (1-p)^{N-i}$$

g. Poisson

Let m = rate parameter or mean = lambda
X = number of occurrences

$$P(X = N) = e^{-m} \frac{m^N}{N!}$$
$$N = 0, 1, \ldots$$

and

$$P = P(X \geqslant N) = e^{-m} \sum_{i=N}^{\infty} \frac{m^i}{i!}$$

h. Hypergeometric

Let N = number of items in a lot
M = sample size
K = number of defective items in the lot
X = number of defective items in the sample
then P (exactly x defectives are in the sample)

$$= P(X = x) = \frac{\binom{K}{x}\binom{N-K}{M-X}}{\binom{N}{M}}$$

$$x = 0, 1, \ldots, M$$

and

$$P = P(X \geqslant x) = \sum_{i=x}^{\min(M,K)} P(X = i)$$

References

Normal

Abramowitz, M. and Stegun, I.A., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, N.B.S. Applied Math. Series No. 55, Government Printing Office, Washington, D.C.

Hastings, C., *Approximations for Digital Computers*, Princeton University Press, Princeton, New Jersey, 1955.

t and *F* Distribution

Aroian. L.A., "Continued fractions for the incomplete beta function", *Annals of Mathematical Statistics,* 12, 1941.

Wall, H.S., *Analytic Theory of Continued Fractions*, D. Van Norstrand, New York, 1948.

Binomial

Aroian. L.A., "Continued fractions for the incomplete beta function", *Annals of Mathematical Statistics,* 12, 1941.

Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons Inc., New York, 1957.

Wall, H.S., *Analytic Theory of Continued Fractions*, D. Van Norstrand, New York, 1948.

Chi-Square

Peizer, D. and Pratt, J., "A normal approximation for binomial, F, Beta and other common related tail probabilities, I", *Journal of the American Statistical Association*, 63, 1968.

Wall, H.S., *Analytic Theory of Continued Fractions*, D. Van Norstrand, New York, 1948.

Poisson and Weibull

Johnson, N. and Kotz, S., *Continuous Univariate Distributions*, Vol. 1 and 2, Houghton-Mifflin, New York.

Hypergeometric

Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons Inc., New York, 1957.

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "DISTR"`
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and `SELECT DISTRIBUTION` is displayed:
   a. Press: KEY#1 (`NORMAL`), to select the normal distribution.
   b. Go to step 5.
   OR:
   a. Press: KEY #2 (`F`), to select the central F distribution.
   b. Go to step 6.
   OR:
   a. Press: KEY #3 (`t`), to select the student's t distribution.
   b. Go to step 7.
   OR:
   a. Press: KEY #4 (`BINOMIAL`), to select the binomial distribution.
   b. Go to step 8.
   OR:
   a. Press: KEY #5 (`POISSON`), to select the Poisson distribution.
   b. Go to step 9.
   OR:
   a. Press: KEY #6 (`CHI-SQ`), to select the chi-square distribution.
   b. Go to step 10.
   OR:
   a. Press: KEY #7 (`WEIBULL`), to select the Weibull distribution.
   b. Go to step 11.
   OR:
   a. Press: KEY #8 (`HYPERGEO`), to select the hypergeometric distribution.
   b. Go to step 12.
5. When `X or X, μ, σ?` is displayed:
   a. Enter: The value of X for which a right tailed probability is desired if $\mu = 0$ and $\sigma = 1$.
   b. Press: (END LINE)
   c. The probability is printed and the program returns to step 4.
   OR:
   a. Enter: The value of X, the mean, μ, and the standard deviation, σ, separated by commas.
   b. Press: (END LINE)
   c. The probability is printed and the program returns to step 4.

   Note: You must enter values for μ and σ and not expressions.
6. When `N, D, X?` is displayed:
   a. Enter: N, the numerator degrees of freedom, D, the denominator degrees of freedom, and X, the value of F for which a right tailed probability is desired. Separate values with a comma.
   b. Press: (END LINE)
   c. The probability is printed and the program returns to step 4.
7. When `N, t?` is displayed:
   a. Enter: N, the degrees of freedom, and t, the t value.

b. Press: END LINE

c. The probability is printed and the program returns to step 4.

8. When N, R, P? is displayed:

a. Enter: N, the number of trials, R, the number of successes, and P, the probability of success at each trial.

b. Press: END LINE

c. The probability is printed and the program returns to step 4.

9. When λ, N? is displayed:

a. Enter: λ, the rate parameter or mean, and N, the number of occurrences.

b. Press: END LINE

c. The probability is printed and the program returns to step 4.

10. When N, X? is displayed:

a. Enter: N, the degrees of freedom, and X, the χ^2 value.

b. Press: END LINE

c. The probability is printed and the program returns to step 4.

11. When α, β, X? is displayed:

a. Enter: The Weibull parameters, α and β, and X, the value for which a right tailed probability is desired.

b. Press: END LINE

c. The probability is printed and the program returns to step 4.

12. When N, M, K, X? is displayed:

a. Enter: N, the number of items in a lot, M, the sample size, K, the number of defective items in the lot, and X, the number of defective items in the sample.

b. Press: END LINE

c. The probability is printed and the program returns to step 4.

Example:

The following results were obtained for the distribution functions with the associated input values.

Normal Prob

```
NORMAL PROBABILITY
PROB(N> 1.2 )= .115069731583

NORMAL PROBABILITY
PROB(N> 5 )= 2.87104999504E-7

NORMAL PROBABILITY
PROB(N> .56 )= .287739682034
```

F Prob

```
F PROBABILITY
N =                    2
D =                    2
PROB(F> 19 )= 5.00000000052E-2

F PROBABILITY
N =                    100
D =                    100
PROB(F> 1 )= 4.9999999913

F PROBABILITY
N =                    4
D =                    1250
PROB(F> 1.2 )= .309025276796
```

t Prob

```
t PROBABILITY
N =                     2
P(t> 12 ) = 3.43646683897E-3

t PROBABILITY
N =                     25
P(t> 2.01 ) = 2.76695014404E-2

t PROBABILITY
N =                     25
P(t> 1.96 ) = .030616282898
```

Right Tail Binomial

```
RIGHT TAIL BINOMIAL PROBABILITY
N =                     15
R =                     12
P =                     .5
P = 1.75781250001E-2

RIGHT TAIL BINOMIAL PROBABILITY
N =                     50
R =                     10
P =                     .4
P = .999242703417

RIGHT TAIL BINOMIAL PROBABILITY
N =                     50
R =                     40
P =                     .4
P = 8.92848366168E-9
```

Right Tail Poisson

```
RIGHT TAIL POISSON PROBABILITY
λ =                     4
N =                     2
P = .908421805547

RIGHT TAIL POISSON PROBABILITY
λ =                     25
N =                     6
P = .999998602885

RIGHT TAIL POISSON PROBABILITY
λ =                     25
N =                     44
P = .000368952614
```

Chi-Square Prob

```
CHI-SQUARE PROBABILITY
N =                    11
P(X> 19 ) = 6.10935302597E-2

CHI-SQUARE PROBABILITY
N =                    100
P(X> 120 ) = 8.44038897858E-2

CHI-SQUARE PROBABILITY
N =                    250
P(X> 275 ) = .132992812576
```

Weibull Prob

```
WEIBULL PROBABILITY
α=                     2
β=                     2
P(X> 1.25 )= 4.39369336234E-2
WEIBULL PROBABILITY
α=                     2
β=                     6
P(X> .5 )= .969233234476
WEIBULL PROBABILITY
α=                     2
β=                     8
P(X> .77 )= .781025672542
```

Right Tail Hypergeometric

```
RIGHT TAIL HYPERGEOMETRIC
       PROBABILITY
N =                    125
M =                    25
K =                    12
X =                    4
P = .196209616814
RIGHT TAIL HYPERGEOMETRIC
       PROBABILITY
N =                    125
M =                    25
K =                    12
X =                    3
P =  .444574089246
RIGHT TAIL HYPERGEOMETRIC
       PROBABILITY
N =                    500
M =                    260
K =                    100
X =                    50
P =  .712215987844
```

Multiple Linear Regression

This program calculates a least squares regression on up to twelve independent variables, including means and variances for all variables. In addition, the program calculates the correlation matrix, a complete analysis of variance table, estimates of variances and t values for the regression coefficients, the multiple correlation coefficient and y − *hat* ($\hat{y}$) evaluation. The equation is of the form: $y = b_0 + b_1x_1 + \ldots + b_kx_k$.

Equations and formulae:

The method employed is to solve the n simultaneous linear normal equations:

$$
\begin{aligned}
c_{11}b_1 + c_{12}b_2 + \ldots + c_{1n}b_n &= c_{1y} \\
c_{21}b_1 + c_{22}b_2 + \ldots + c_{2n}b_n &= c_{2y} \\
&\vdots \\
c_{n1}b_1 + c_{n2}b_2 + \ldots + c_{nn}b_n &= c_{ny}
\end{aligned}
$$

where

$$c_{ij} = \Sigma x_i x_j - \frac{\Sigma x_i \, \Sigma x_j}{n}$$

$$c_{iy} = \Sigma x_i y - \frac{\Sigma x_i \, \Sigma y}{n}$$

$$c_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$$

and

$$b_0 = \bar{y} - \sum_i b_i \bar{x}_i$$

By inverting the $n \times n$ matrix $C = (c_{ij})$ to obtain $C^{-1} = (c_{ij}')$, we then have $b_i = \sum_j c_{ij}' \, c_{jy}$.

The multiple correlation coefficient is given by

$$R^2 = \sum_{j=1}^{n} \left(b_j \frac{s_{jy}^2}{s_y^2} \right)$$

where the variance of y

$$S_y^2 = \frac{1}{n-1}\Sigma(x_{1y} - y)^2$$

and the covariance

$$S_{1y}^2 = \frac{1}{n-1}\Sigma((x_{1j} - x_1)(x_{1y} - y))$$

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: `LOAD "MLR"`
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and `SELECT OPTION` is displayed:
   a. Press: **KEY #5** (`HELP`), if you need a more detailed explanation.
   b. After the explanation is displayed, go to step 4.
   OR:
   a. Press: **KEY #1** (`ENTER`), to enter data.
   b. Go to step 5.
   OR:
   a. Press: **KEY #2** (`PRINT`), to print the data in a file on tape.
   b. Go to step 21.
   OR:
   a. Press: **KEY #6** (`EDIT`), to edit a data set on tape.
   b. Go to step 24.
   OR:
   a. Go to step 19, if data has already been entered.
5. When `MAX. INDEPENDENT VAR. = 12:K=?` is displayed:
   a. Enter The number of independent variables.
   b. Press: (END LINE)
   Note: If the number of independent variables is not in the range of 2 to 12, the program will beep and go to step 5.
6. When `# OF DATA SETS?` is displayed:
   a. Enter: The number of data sets.
   b. Press: (END LINE)
7. When `ENTER FROM KEYBOARD/TAPE:K/T?` is displayed:
   a. Enter: K, if the data is to be entered from the keyboard.
   b. Press: (END LINE)
   c. Go to step 9.
   OR:
   a. Enter: T, if the data is to be entered from the tape.
   b. Press: (END LINE)
   c. Go to step 8.
   Note: You must enter either K or T or the program will beep and go to step 7.
8. When `ENTER FILE NAME?` is displayed:
   a. Enter: The name of the file.
   b. Press: (END LINE)
   c. After the data sets have been read and printed, go to step 19.

REFERENCES:.

1. Graybill, F. A., *An Introduction to Linear Statistical Models, Vol. I.*, McGraw-Hill, 1961.
2. Yamane, T., *Statistics, An Introductory Analysis*, Harper and Row, 1967.

9. When STORE DATA SETS AS ENTERED: Y/N? is displayed:
   a. Enter: Y, if the data is to be stored as entered.
   b. Press: END LINE
   c. Go to step 10.

   OR:

   a. Enter: N, if the data is not to be stored.
   b. Press: END LINE
   c. Go to step 12.

   Note: You must enter either Y or N or the program will beep and go to step 9.
10. When ENTER FILE NAME? is displayed:
    a. Enter: The file name.
    b. Press: END LINE
11. When CREATE FILE: Y/N? is displayed:
    a. Enter: Y, to create the file.
    b. Press: END LINE

    OR:

    a. Enter: N, if the file already exists.
    b. Press: END LINE

    Note: You must enter either Y or N or the program will beep and go to step 11.
12. When Y=? is displayed:
    a. Enter: The value of the dependent variable.
    b. Press: END LINE
13. When X(__)=? is displayed:
    a. Enter: The value of the specified independent variable.
    b. Press: END LINE
14. Repeat step 13 for each independent variable.
15. When EDIT DATA SET: Y/N? is displayed:
    a. Enter: Y, to edit the data set just entered.
    b. Press: END LINE
    c. Go to step 16.

    OR:

    a. Enter: N, if the data set is correct.
    b. Press: END LINE
    c. Go to step 18.

    Note: You must enter either Y or N or the program will beep and go to step 15.
16. When ENTER INDEX TO CHANGE? is displayed:
    a. Enter: The index of the value to change.
    b. Press: END LINE

    Note: If the index is not valid, the program will beep and go to step 16.
17. When NEW VALUE=? is displayed:
    a. Enter: The new value.
    b. Press: END LINE
    c. Go to step 15.
18. Repeat steps 12 through 17 for each data set.
19. When DONE is displayed:
    a. Press: **KEY #3** (MLR), to perform the regression.
    b. After the regression is done, go to step 20.

    Note: If MORE DATA SETS NEEDED or MUST ENTER DATA FIRST! is displayed, go to step 4 and enter at least K+1 data sets before trying this option.

    OR:

    a. Go to step 4 to enter, edit, or print data.
20. When DONE is displayed:
    a. Press: **KEY #4**(Y EST.), to estimate a value of Y based on the regression coefficients.
    b. Go to step 38.

    OR:

    a. Go to step 4 to enter, edit, or print data.
21. When ENTER FILE NAME TO PRINT? is displayed:
    a. Enter: The name of the file to print.
    b. Press: END LINE

    Note: If ERROR ON ASSIGNMENT is displayed, go to step 21 and enter a name of a file on the tape.

22. When MAX. INDEPENDENT VAR.= 12:K=? is displayed:
    a. Enter: The number of independent variables.
    b. Press: END LINE

    Note: If the number of independent variables is not in the range of 2 to 12, the program will beep and go to step 22.

23. When ENTER # DATA SETS IN __? is displayed:
    a. Enter: The number of data sets in the file.
    b. Press: END LINE
    c. After the data sets are printed, go to step 4.

24. When ENTER FILE NAME? is displayed:
    a. Enter: The name of the file to edit.
    b. Press: END LINE

    Note: If ERROR ON ASSIGNMENT is displayed, go to step 24 and enter the name of a file on the tape.

25. When MAX INDEPENDENT VAR.= 12:K=? is displayed:
    a. Enter: The number of independent variables.
    b. Press: END LINE

    Note: If the number of independent variables is not in the range of 2 to 12, the program will beep and go to step 25.

26. When ENTER # DATA SETS IN __? is displayed:
    a. Enter: The number of data sets in the file.
    b. Press: END LINE

27. When EDIT CODE? is displayed:
    a. Enter: 0, if the edit is complete.
    b. Press: END LINE

    OR:

    a. Enter: 1, to correct a value in a data set.
    b. Press: END LINE

    OR:

    a. Enter: 2, to delete a data set.
    b. Press: END LINE
    c. Go to step 33.

    OR:

    a. Enter: 3, to add a data set.
    b. Press: END LINE
    c. Go to step 34.

    Note: If the code is not in the range of 0 to 3, the program will beep and go to step 27.

28. When ENTER DATA SET # TO CORRECT? is displayed:
    a. Enter: The index of the data set to correct.
    b. Press: END LINE

    Note: If the index is greater than the number of data sets in the file, the program will beep and go to step 28.

29. When Y=? is displayed:
    a. Enter: The value of the dependent variable.
    b. Press: END LINE

30. When X(__)=? is displayed:
    a. Enter: The value of the specified independent variable.
    b. Press: END LINE

31. Repeat step 30 for each independent variable.
32. Go to step 27.
33. When ENTER DATA SET # TO DELETE? is displayed:
    a. Enter: The index of the data set to delete.
    b. Press: END LINE
    c. After the number of remaining data sets is printed, go to step 27.

    Note: If the index is greater than the number of data sets in the file, the program will beep and go to step 33.

34. When Y=? is displayed:
    a. Enter: The value of the dependent variable.
    b. Press: END LINE

35. When X(__)? is displayed:
    a. Enter: The value of the specified independent variable.

b. Press: (END LINE)

36. Repeat step 35 for each independent variable.
37. Go to step 27.
38. When X(___)=? is displayed:

    a. Enter: The value of the specified independent variable.

    b. Press: (END LINE)

39. Repeat step 38 for each independent variable.
40. After the estimated value of Y is printed, go to step 20.

Example:

Using the following data with 3 independent variables and 6 data sets, perform a multiple linear regression.

Y	X(1)	X(2)	X(3)
3.0	5.0	1.0	8.0
4.2	2.0	1.5	9.0
4.5	1.0	1.6	1.0
6.0	7.0	2.0	4.0
6.5	9.0	2.1	12.0
8.4	4.0	2.5	2.0

```
MULTIPLE LINEAR REGRESSION

DATA SET  1:  I        X(I)
              1       5.0000
              2       1.0000
              3       8.0000
     (Y)      4       3.0000

DATA SET  2:  I        X(I)
              1       2.0000
              2       1.5000
              3       9.0000
     (Y)      4       4.2000

DATA SET  3:  I        X(I)
              1       1.0000
              2       1.6000
              3       1.0000
     (Y)      4       4.5000

DATA SET  4:  I        X(I)
              1       7.0000
              2       2.0000
              3       4.0000
     (Y)      4       6.0000

DATA SET  5:  I        X(I)
              1       9.0000
              2       2.1000
              3      12.0000
     (Y)      4       6.5000
```

```
DATA SET  6:  I        X(I)
              1         4.0000
              2         2.5000
              3         2.0000
     (Y)      4         8.4000

VAR          MEAN      VARIANCE
X(  1)        4.6667     9.0667
X(  2)        1.7833      .2777
X(  3)        6.0000    18.8000
X(  4)        5.4333     3.7147

CORRELATION MATRIX

   1     1.000

   2      .324     1.000

   3      .536     -.254     1.000

   4      .340      .987     -.251
           1.000

MULTIPLE CORRELATION =     .976

ANALYSIS OF VARIANCE

SOURCE/DF     SS       MS      F
TOTAL    5    18.6
REG      3    18.1      6.0   26.7
 X(  1) 1      2.1      2.1    9.5
 X(  2) 1     16.0     16.0   70.5
 X(  3) 1      0.0      0.0    0.0
RESID    2     0.5      0.2


COEFFICIENTS
  I      B(I)    VARIANCE    TVALUE

  0    -0.918
  1     0.028    0.010      0.273
  2     3.533    0.255      6.991
  3    -0.013    0.005     -0.189
```

Appendix

The documentation for the programs contained in the General Statistics Pac is available by executing the comments program, STCOM. By executing this program, you can obtain the comments for any or all of the programs in the pac. The definitions of the major variables used in each program are also available.

User Instructions

1. Insert the General Statistics Pac cartridge into the tape transport.
2. To load the program:
   a. Type: LOAD "STCOM"
   b. Press: (END LINE)
3. To start the program:
   a. Press: (RUN)
4. When the keys are labelled and SELECT PROGRAM is displayed:
   a. Press: KEY #1 (ONESAM) to select comments for one sample analysis programs (ONESAM and ONESM2).
   b. Go to step 6.

   OR:

   a. Press: KEY #2 (PAIRED) to select comments for paired sample analysis programs (PAIRED, PAIR2, and PAIR3).
   b. Go to step 8.

   OR:

   a. Press: KEY #3 (TESTS) to select comments for test statistics programs (CHISQR, RXCCHI, TSTAT2, ONEAOV, and TWOAOV).
   b. Go to step 10.

   OR:

   a. Press: KEY #4 (DISTR) to obtain the comments for the distributions program (DISTR).
   b. Go to step 5.

   OR:

   a. Press: KEY #5 (MLR) to obtain the comments for the multiple linear regression program (MLR).
   b. Go to step 5.

   OR:

   a. Press: KEY #6 (DONE).
   b. The program will stop.
5. When VARIABLE DEFINITIONS: Y/N? is displayed:
   a. Enter: Y, to obtain the variable definitions.
   b. Press: (END LINE)
   c. Go to step 4.

   OR:

   a. Enter: N, if the variable definitions are not wanted.
   b. Press: (END LINE)
   c. Go to step 4.
6. When the keys are relabelled and SELECT OPTION is displayed:
   a. Press: KEY #1 (ONESAM) to obtain the comments for the program ONESAM.
   b. Go to step 7.

   OR:

a. Press: **KEY #2** (ONESM2) to obtain the comments for the program ONESM2.
b. Go to step 7.

OR:

a. Press: **KEY #3** (DONE).
b. Go to step 4.

7. When VARIABLE DEFINITIONS: Y/N? is displayed:
   a. Enter: **Y**, to obtain the variable definitions.
   b. Press: (END LINE)
   c. Go to step 6.

   OR:

   a. Enter: **N**, if the variable definitions are not wanted.
   b. Press: (END LINE)
   c. Go to step 6.

8. When the keys are relabelled and SELECT OPTION is displayed:
   a. Press: **KEY #1** (PAIRED) to obtain the comments for the program PAIRED.
   b. Go to step 9.

   OR:

   a. Press: **KEY #2** (PAIR2) to obtain the comments for the program PAIR2.
   b. Go to step 9.

   OR:

   a. Press: **KEY #3** (PAIR3) to obtain the comments for the program PAIR3.
   b. Go to step 9.

   OR:

   a. Press: **KEY #4** (DONE).
   b. Go to step 4.

9. When VARIABLE DEFINITIONS: Y/N? is displayed:.
   a. Enter: **Y**, to obtain the variable definitions.
   b. Press: (END LINE)
   c. Go to step 8.

   OR:

   a. Enter: **N**, if the variable definitions are not wanted.
   b. Press: (END LINE)
   c. Go to step 8.

10. When the keys are relabelled and SELECT PROGRAM is displayed:
    a. Press: **KEY #1** (CHISQR) to obtain the comments for the program CHISQR.
    b. Go to step 11.

    OR:

    a. Press: **KEY #2** (RXCCHI) to obtain the comments for the program RXCCHI.
    b. Go to step 11.

    OR:

    a. Press: **KEY #3** (TSTAT2) to obtain the comments for the program TSTAT2.
    b. Go to step 11.

    OR:

    a. Press: **KEY #4** (ONEAOV) to obtain the comments for the program ONEAOV.
    b. Go to step 11.

    OR:

    a. Press: **KEY #5** (TWOAOV) to obtain the comments for the program TWOAOV.
    b. Go to step 11.

    OR:

    a. Press: **KEY #6** (DONE).
    b. Go to step 4.

11. When VARIABLE DEFINITIONS: Y/N? is displayed:
    a. Enter: **Y**, to obtain the variable definitions.
    b. Press: (END LINE)
    c. Go to step 10.

    OR:

    a. Enter: **N**, if the variable definitions are not wanted.
    b. Press: (END LINE)
    c. Go to step 10.

HEWLETT
PACKARD

For additional information please contact the nearest authorized HP-85 dealer or your local Hewlett-Packard sales office.

Printed in U.S.A.