hp calculators

HP 17bII+ Statistics – Non-linear Regression

Statistics on the HP 17bII+

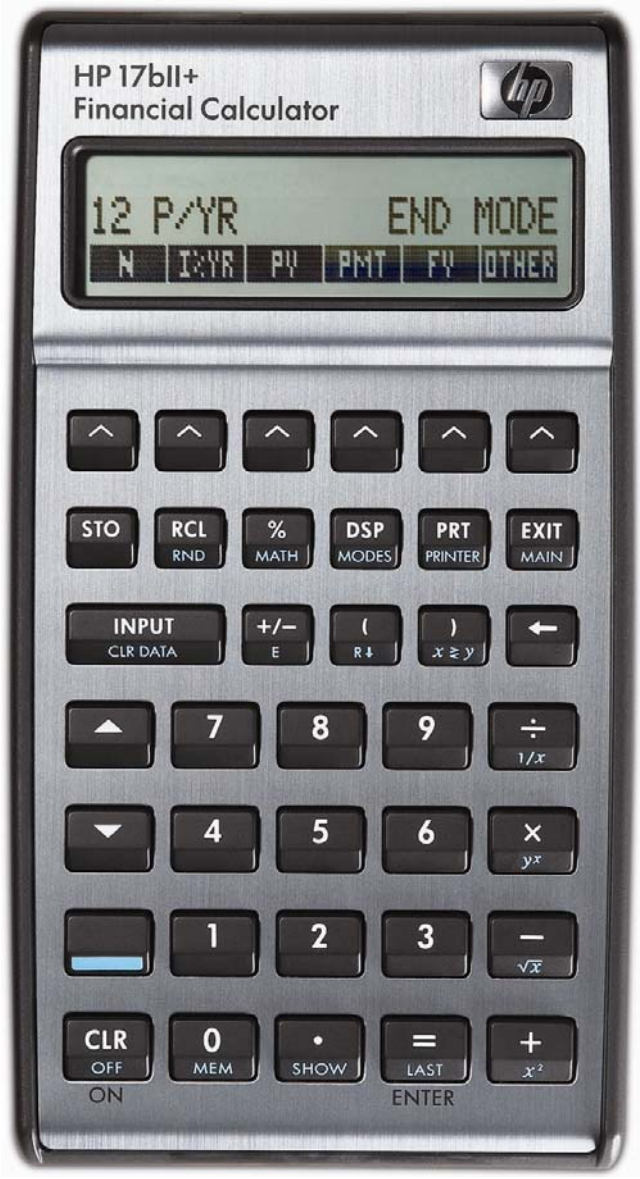
Non-linear regression

Sum lists

Entering data values

Non-linear regression on the HP 17bII+

Practice solving problems involving Non-linear regression



Statistics on the HP 17bII+

The HP 17bII+ has many built-in statistics functions that apply to finding averages and standard deviations as well as regression and correlation.

Non-linear regression

Linear regression finds the equation of a line in the form Y = mX + b, where m is the slope of the line and b is the Y-intercept, that "best fits" a set of X and Y data. Linear regression calculates the equation for this line by minimizing the sum of the squared residuals between the actual data points and the predicted data points using the estimated line's slope and intercept.

Non-linear regression does the same thing, but uses an equation that is not a line. These equations can come in many forms: power curves, exponential curves, or logarithmic curves, to nave a few. In these curves, m and b are still calculated, but have slightly different meanings, depending on the curve involved, as shown below.

Logarithmic regression:	$y = B + M\ LN(x)$
Exponential regression:	$y = B\ e^{Mx}$
Power regression:	$y = B\ x^{M}$

Sum lists

On the HP 17bII+, statistics problems are solved in the Sum environment. This is entered from the main menu by pressing SUM.

Figure 1

Figure 2

On this screen, the menu keys on the bottom of the display perform these functions:

CALC – Ends the entry of statistical data into the sum list and enters the statistics calculation environment.

TOTAL – Displays the total of the data values in the current sum list.

The other menu functions, INSR, DELET, NAME, and GET allow you to edit and save a sum list. They are covered in another learning module.

The HP 17bII+ uses a different approach to handling statistics problems than other HP calculators. Rather than storing the data values in registers, the HP 17bII+ stores data values in lists that can be named and saved for future use. This also provides the flexibility to have multiple lists of statistical data stored within the calculator at the same time, limited only by the available calculator memory.

Entering data values

The initial data value is keyed and entered using the INPUT CLR DATA key. The remaining data values are entered in the same manner. Data values can be keyed or computed. They are added to the list when you press INPUT CLR DATA = LAST. Note to RPN users: This is an instance where there is a difference between the INPUT CLR DATA and = LAST keys. If you wish to compute a value and add it to a sum list, use = LAST rather than INPUT CLR DATA to serve as ENTER. Pressing INPUT CLR DATA in RPN mode while entering data into a sum list will add the number displayed to the list.

When you have entered all data values, press EXIT followed by CALC to enter the calculation environment, which appears as shown below in figure 3. The menu choices of TOTAL, MEAN, MEDN, STDEV and RANG are covered in other learning modules. Pressing MORE shows the second page of the statistics calculation environment as shown in figure 4 below.

Figure 3

Figure 4

MIN, MAX, and SORT are covered in other learning modules. FRCST displays a series of menus for calculations involving two variables for curve fitting or statistical analysis requiring two lists.

Non-linear regression on the HP 17bII+

On the HP 17bII+, linear regression problems involve entering two sum lists – one for the dependent variable (X) and one for the independent variable (Y). These must be entered and saved with a name in order to associate them together for a regression problem. Once these have been entered, pressing FRCST will allow you to choose a list for the X values and a list for the Y values. Once chosen, the screen displays the forecasting menu as shown below. The first two menu positions represent the independent and dependent variables chosen as you entered the FRCST environment.

Figure 5

X – This location represents the independent variable. Enter a value for the dependent variable and press this key to estimate the corresponding value of the independent variable. The name shown at this menu position will depend upon the list chosen. In the rest of the statistics options, this is referred to as the X-variable list.

SALES – This location represents the dependent variable. Enter a value for the independent variable and press this key to estimate the corresponding value of the dependent variable. The name shown at this menu position will depend upon the list chosen. In the rest of the statistics options, this is referred to as the Y-variable list.

CORR – Computes and displays correlation coefficient of the independent and dependent variables. This value will be between -1 and +1. The better the "fit" of the model to the X and Y variable lists, the closer this value will be to -1 or +1. Two lists that have no relationship to each other at all would have a correlation of 0.

M – Computes and displays M. This is the slope for the linear regression model.

B – Computes and displays B. This is the y-intercept for the linear regression model.

MORE – Displays the second page of forecasting menu options as shown below in figure 6.

Figure 6

MODL – Displays a screen of choices for regression models to be used. The HP 17bII+ provides a choice of four models. Linear LIN, Logarithmic LOG, Exponential EXP and Power PWR fits.

W.MN
G.SD
SIZE – These three functions are covered in another learning module.

MORE – Displays the third page of forecasting menu options, the functions of which are discussed in another module.

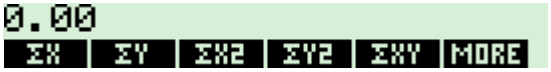


Figure 7

Practice solving non-linear regression problems

Example 1: John's store has had sales for the last 5 months of $150, $165, $180, $195, and $220. Use a power regression to predict sales for months 6 and 7 and also predict when estimated sales would reach $250. What is the correlation for the regression?

Solution: [ON] [EXIT] [STAT]

If you already have at least one sum list, press: [CLEAR] [ENTER]

[1] [INPUT] [2] [INPUT] [3] [INPUT] [4] [INPUT] [5] [INPUT] [EXIT]

[NAME] [RCL] [X] [INPUT]

[CLEAR] [ENTER]

[150] [MEM] [INPUT] [165] [INPUT] [180] [MEM] [INPUT] [195] [INPUT] [220] [MEM] [INPUT] [EXIT]

[NAME] [RCL] [5] [RCL] [X] [LN] [LN] [RCL] [Y] [RCL] [5] [INPUT]

[CALC] [MORE] [FAST]

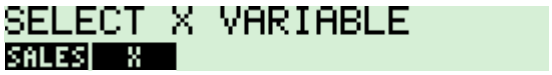


Figure 8

[X]

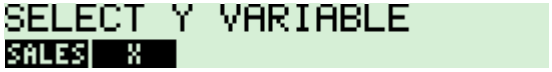


Figure 9

[SALES]

Figure 10

Since the regression model choice shown is linear and we need a power curve model, press:
[MORE] [MOD]

Figure 11

[PWR]

Figure 12

[M]

Figure 13

[B]

Figure 14

This means the power curve equation fit to the data is $y = 145.51\ x^{0.22}$

[6] [X] [SALES]

Figure 15

[7] [X] [SALES]

Figure 16

[2] [5] [0 MEM] [SALES] [X]

Figure 17

[CORR]

Figure 18

Answer: Sales in month 6 are predicted to be $217 and in month 7 $225. Sales are projected to reach $250 at month 11 or 12. The correlation is 0.96, which indicates a fairly strong relationship and predictive ability.

Example 2: Assuming the data from example 1 is still in the HP 17bII+ and that you have just finished working example 1, use the other 3 forecasting models and check on the correlation coefficient. Which of the four models has the highest correlation coefficient?

Solution: The correlation for the power model was already determined to be 0.96 in example 1. To compute the correlation for the linear model, press [MORE] until you see the [MODL] menu label.

[MODL]

Figure 19

LIN CORR

Figure 20

MORE MORE LOG CORR

Figure 21

MORE MORE EXP CORR

Figure 22

Answer: The results indicated a linear model correlation of 0.99, a logarithmic model correlation of 0.95, and an exponential model correlation of 1 (actually, 0.9977). Both the linear model and exponential model have a higher correlation than the power curve from example 1.