hp calculators

HP 10BII Statistics – Linear Regression

Statistics on the HP 10BII

Linear regression and trend lines

Linear regression on the HP 10BII

Practice solving problems involving linear regression

Statistics on the HP 10BII

The HP 10BII has many built-in statistics functions that apply to finding averages and standard deviations as well as linear regression, correlation and rearranging items.

Linear regression and trend lines

A trend line is actually an equation of a line in the form $Y = bX + a$, where b is the slope of the line and a is the Y-intercept. 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. Once the slope and intercept have been calculated, it is fairly easy to substitute other values for X and predict a corresponding value for Y, or to substitute a value for Y and predict a value for X. When the X value is a measure of time (months or years, for example), the equation is specifically referred to as a trend line. These are often used to predict future sales growth given past sales data. Be aware, however, that it is rarely a good idea to use such an equation to predict too far into the future from the actual data used, since circumstances can change rather quickly. Also be aware that these predictions are linear in nature and make no adjustment for any seasonality that may exist.

Linear regression is also often used to estimate the fixed and variable components from a company's or department's total costs. In these circumstances, the values for X are usually the cost driver for the organization or department. Examples might include units produced, hours worked, hours of machine time, and others. The values for Y are the total cost for that level of X input. The computed slope of the linear regression line will indicate the variable cost per unit of X, while the computed Y-intercept will indicate the fixed cost.

In many or most circumstances, this type of cost analysis will generate slopes and Y-intercepts that make sense in the real world. It is sometimes possible, though, that the fixed cost component in particular may not make any sense. The generated Y-intercept (fixed cost) might be negative, for example, to make the linear regression line fit the observed cost data as closely as possible. Be aware, as well, that it is rarely a good idea to use such an equation to predict too far into the future from the actual data used, since circumstances can change rather quickly. In other words, if you fit a line using cost data for units produced from 500 to 1500 a month, making cost predictions using forecasted production levels of 5000 units a month may generate unreliable results. Also, since time is not a variable in these calculations, the order in which the costs are input as data points does not matter – you may enter the data points in any order desired.

Linear regression on the HP 10BII

Before beginning a new statistics problem, it is always a good idea to clear the statistics registers by pressing the yellow-shifted function of the →M key, or ▬ cLΣ .

On the HP 10BII, linear regression problems involve entering pairs of numbers by keying the first number, pressing INPUT and then entering the second number and pressing the Σ+ key. Note that each time the Σ+ key is pressed, the HP 10BII will display a number representing the number of data pairs entered into the statistics registers. Wrong pairs of numbers can be deleted by entering the incorrect pair of numbers again and pressing ▬ Σ- . Once all the pairs of numbers have been entered, pressing C ▬ ŷm will display the y-intercept. Pressing ▬ SWAP will display the slope.

Once the linear regression equation has been computed, it is possible to enter an X value and have the HP 10BII predict the corresponding Y value using the equation. This is done by entering the X value and pressing ▬ ŷm . It is also possible to enter a Y value and have the HP 10BII predict the corresponding X value. This is done by entering the Y value and pressing ▬ x̂r .

Practice solving linear regression problems

<u>Example 1:</u>   John's store has had sales for the last 5 months of $150, $165, $160, $175, and $170. Use a trend line to predict sales for months 6 and 7 and also predict when estimated sales would reach $200.

<u>Solution:</u>

| ▬ | CLΣ |

1 INPUT 1 5 0 Σ+ 2 INPUT 1 6 5 Σ+ 3 INPUT 1 6 0 Σ+
4 INPUT 1 7 5 Σ+ 5 INPUT 1 7 0
C ▬ ŷ,m
▬ SWAP
6 ▬ ŷ,m
7 ▬ ŷ,m
2 0 0 ▬ x̂,r .

<u>Answer:</u>   The slope is 5 and the y-intercept is 149. The linear regression equation is therefore Y = 5X + 149. Sales in month 6 are predicted to be $179 and in month 7 are expected to be $184. Sales are predicted to reach $200 between months 10 and 11 (actual answer is 10.2).

<u>Example 2:</u>   Johnson's Chair Company has experienced the following costs for the first 6 months of the year:

# Chairs Made	Total Costs
5,000	$120,000
5,500	$122,100
4,800	$118,540
5,300	$122,400
4,950	$119,100
5,150	$124,200

What estimate would a linear regression equation produce for Johnson's fixed and variable cost? What are the total costs predicted if 5,400 chairs were to be made? If the total costs were $125,000, how many chairs would you estimate had been produced?

<u>Solution:</u>   The X values will be the number of chairs produced. The Y values will be the total costs.

▬ CLΣ
5 0 0 0 INPUT 1 2 0 0 0 0 Σ+
5 5 0 0 INPUT 1 2 2 1 0 0 Σ+
4 8 0 0 INPUT 1 1 8 5 4 0 Σ+
5 3 0 0 INPUT 1 2 2 4 0 0 Σ+
4 9 5 0 INPUT 1 1 9 1 0 0 Σ+
5 1 5 0 INPUT 1 2 4 2 0 0 Σ+
C ▬ ŷ,m
▬ SWAP
5 4 0 0 ▬ ŷ,m
1 2 5 0 0 0 ▬ x̂,r

<u>Answer:</u>   The linear regression equation generated is of the form: Y = 6.18X + 89449.38. The slope of 6.18 is the estimate for the variable cost and the Y-intercept of 89,449.38 is the estimate for the fixed cost. The total cost estimate if 5,400

chairs were made is $122,806. The estimated number of chairs made if the total costs were $125,000 is 5,755 chairs.

Example 3: The stamping department cost analyst is reviewing the total cost compared with the number of machine hours used for the last 4 weeks.

# Machine Hours	Total Costs
350	$55,000
375	$57,300
400	$58,100
360	$56,250

What estimate would a linear regression equation produce for the stamping department's fixed and variable cost? What are the total costs predicted if 380 machine hours are used next week? If the total costs were $60,000, how many machine hours would you estimate had been used?

Solution: The X values will be the number of machine hours used. The Y values will be the total costs.
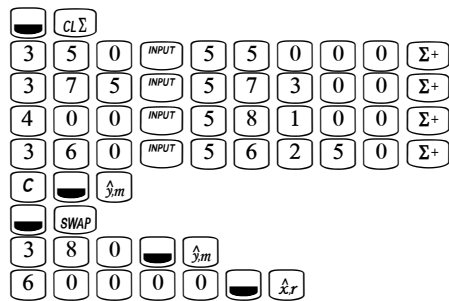
Answer: The linear regression equation generated is of the form: Y = 58.99X + 34,763.66. The slope of 58.99 is the estimate for the variable cost per machine hour and the Y-intercept of 34,763.66 is the estimate for the fixed cost. The total cost estimate if 380 machine hours are used is $57,178. The estimated number of machine hours used if the total costs were $60,000 is 427.8 machine hours.