Sinusoidal Regression

Prerequisite Skills

This lesson requires the use of the following skills:

- plotting points on the coordinate plane
- evaluating sine functions
- graphing sine functions
- analyzing the graphs of functions

Introduction

Many things happen in cycles. For example, in temperate zones, the number of hours of daylight will increase steadily through the year until they reach a maximum in the summer, at which point the hours of daylight will begin to decrease. This pattern repeats the next year. Periodic phenomena such as this can be modeled using sine functions.

Key Concepts

- Recall that the process of finding a model to fit a data set is called **regression analysis**. The model fitted to the data is called the regression model, or just the **regression**.
- A sine function can be used to model the data if:
  - the data is curved
  - the data shows periods of increase and decrease
  - the data is periodic; that is, if the data pattern repeats, or could be expected to repeat
- A sine curve that has been fitted to a data set is called a **sinusoidal regression**. Sinusoidal is a shorthand description of the shape of the sine curve.
- In contrast with data with a quadratic trend, sinusoidal data tends to have a flared U-shape. While quadratic models are sometimes used to model a single period of data, their fit tends not to be as good on the flared portions of the curve.
- This tendency can be reflected in the residual plots, yielding a U-shaped graph even though visually the graph appears to be a reasonable fit.
- You can use a calculator or other graphing software to find a sinusoidal regression model for a data set. Note that TI calculators assume the calculator is in radian mode for sinusoidal regression. Be sure to check that your calculator is set to radian mode before graphing calculator-generated sinusoidal regressions.

On a TI-83/84:

Step 1: Press [STAT] to bring up the statistics menu. The first option, 1: Edit, will already be highlighted. Press [ENTER].

Step 2: Arrow up to L1 and press [CLEAR], then [ENTER], to clear the list. Repeat this process to clear L2 and L3, if needed.

Step 3: Enter the ordered pairs in the L1 and L2 lists. Make sure to enter the x-coordinates in L1 and the y-coordinates in L2.

Step 4: Press [STAT]. Arrow to the CALC menu. Press C: SinReg.

Step 5: Press [(], then [2ND][1] to type "L1" for Xlist. Press [,], then [2ND][2] to type "L2" for Ylist. Press [)] to close the parentheses.

Step 6: Press [ENTER] to calculate. The parameter values will appear on the screen.

On a TI-Nspire:

Step 1: Press [home]. Arrow over to the spreadsheet icon, the fourth icon from the left, and press [enter].

Step 2: To clear the lists in your calculator, arrow up to the topmost cell of the table to highlight the entire column, then press [menu]. Choose 3: Data, then 4: Clear Data. Repeat for each column as necessary.

Step 3: Arrow up to the topmost cell of the first column, labeled "A." Press [X] [enter] to type x. Then arrow over to the second column, labeled "B." Press [Y][enter] to type y.

Step 4: Arrow down to cell A1 and enter the first x-value from the ordered pairs. Press [enter]. Enter the second x-value in cell A2 and so on.

Step 5: Move over to cell B1 and enter the first y-value. Press [enter]. Enter the second y-value in cell B2 and so on.

Step 6: To fit an equation to the data points, press [menu] and select 4: Statistics, then 1: Stat Calculations, then C for sinusoidal regression. Select "x" from the X List pop-up menu. Press [tab] to move to the Y List pop-up menu, then select "y" from the options. Tab to "OK" and press [enter]. The parameter values will appear on the screen.

- Note that most graphing software returns sinusoidal regression models in a form equivalent to $y = a \sin(bx + c) + d$. (Some software places the constant d in front of the sine expression.)
- To compare a data set and a function, plot the function on the same coordinate plane as the scatter plot of the data set.

- Graph a sine function by plotting at least five points and drawing a curve through those points.
- Recall that a function is a good fit for the data set if it passes closely to the data points and if some of the data points are above the curve and some are below the curve.
- Recall that a function that is a good fit for the data can be used to make estimates for data not included in the plotted data set.
- Evaluate a function algebraically for a given value of x or y by substituting the given value for x or y and solving for the remaining variable.
- Evaluate a function graphically for a given value of x or y by finding the point on the graph of the function with the known coordinate, then finding the corresponding x- or y-value of that point.
- You can also analyze the **residuals** to assess how well a model fits the data. Recall that a residual is the difference between the function-predicted y-coordinate and the y-coordinate of the actual data point.
- Recall that a **residual plot** is used to visually assess the fit. To create a residual plot, subtract each predicted y-value from the actual y-value for each point in the data set. Then plot the differences against their corresponding x-values.
- A residual plot with a random shape indicates the function is a good fit for the data.
- A residual plot with a U-shape or other distinct pattern indicates that the function is not a good fit for the data.
- If two separate regression models both have similar residual plots, the model whose residuals are closer to 0 is a better fit.

Common Errors/Misconceptions

- using a sine function to model a data set that does not exhibit sinusoidal characteristics
- confusing x and y when graphing data points or analyzing a graph
- confusing parameters in technology-generated regressions
- using radian inputs for a model generated in degree mode or vice-versa