Advanced Techniques in Google Sheets

Stony Brook Physics Labs

Fall 2018

Contents

This document assumes you are already familiar with the basic features of Google Sheets presented in the other documents.

If so, here are a few more advanced features that Google Sheets offers that you can/will make use of in this class.

1   Best Fit Line Calculations

1.1   Summary

In this class, we'll often want to run a line of best fit through some data.

Unfortunately, the default methods that Google Sheets (or Excel) gives for doing this does not give something we want, the uncertainty in the numbers it gives (slope and intercept). So, we'll have to learn how to do it a bit more directly.

To do this, we will use the `LINEST` function. This function makes a line-of-best fit, but has the ability to tell you more than the automatic fitter does.

The results of this function should agree perfectly with the PHY121/122 Plotting Tool. (They will not agree with the PHY133/134 plotting tool, which also takes error bars on your data into account when fitting the line.)

Whenever a best fit line is required in this class, there will be a small side-table for it. You will enter the `LINEST` function into the top-left corner, just

below where the units go.

Note that this is *not* a replacement for actually having plots in your lab report. It is just a good way to get your numbers into your Google Sheet without rounding or re-entering data (which avoids typos and makes it easier to deal with in lab, if desired).

1.2 How to Input

There are four parameters of this function: in the documentation, it is described as LINEST(known_data_y, known_data_x, calculate_b, verbose). Their meaning is as follows:

- known_data_y: This is the data you're using for your y-axis. If this is a range of cells listed vertically (say, B5 through B19), then it will be written as B5:B19. This should be filled in automatically if you highlight a range of cells.

- known_data_x: Data you're using for your x-axis. Same format as the y-data.

- calculate_b: Enter 1 or TRUE if you want an intercept for your plot; enter 0 or FALSE if you want the best fit line to be forced to pass through the origin.

- verbose: This should always be entered with 1 or TRUE. This tells Google Sheets that we want all of the information it can give us and not just the slope/intercept, which is how we get the uncertainties.

Thus, if you want to plot B2 through B9 as your x-data and C2 through C9 as your y-data, with the best-fit line locked through the origin, then you should enter into that upper-left box the command =LINEST(C2:C9,B2:B9,TRUE,TRUE).

1.3 Understanding the Outputs

This will auto-fill not only the cell you enter it into, but ten cells (two wide, five down) with that as the upper-left cell. There is a variety of data this will output; we will only use the first five cells.

The first two cells (the top row) contain the estimates of your parameters: the first cell is the slope, and the second cell is the intercept. The next two cells (the second row) are the uncertainties in those quantities - the third cell is the slope uncertainty and the fourth cell is the intercept uncertainty.

The fifth cell contains the R^2 value. This *correlation coefficient* is a measure of how much your data "looks like a line": if it is close to 1, your data is highly

linear.

The remaining cells contain various statistical information that we will not be using. If you are curious, though, you can find out what they are telling you at `https://support.google.com/docs/answer/3094249?hl=en`.

1.4 LINEST in Excel

In Excel (or LibreOffice or many other programs), you need to do one modification to the above procedure.

After entering the formual, press `Ctrl`+`Shift`+`Enter` (on a Mac, `⌘`+`Enter`) instead of just `Enter`. This tells Excel to fill it out as a *table* formula, rather than just one cell, and will give you all parameters listed above (as opposed to just the slope).