Uncertainties via Monte Carlo Methods

PHYS 310

Physics Department, Bucknell University

I. INTRODUCTION

In PHYS 310 there are several occasions on which you will have to fit some data set to some function and determine the "best-fit parameters." Software packages, like Mathematica, gnuplot, LoggerPro, and Excel (and others) all do some fitting, and they spit out uncertainties in best fit parameters. In these notes I want to talk about what these uncertainties mean and when it's reasonable to use them. Perhaps more important, I want to discuss a "universal" method for determining such uncertainties based on simulating reasonable hypothetical data sets. To introduce this "universal" procedure, I will apply it to a succession of data analysis problems that start with simplest problems, for which there are well-known analytical solutions, and progress to problems that don't have simple analytical solutions. These notes are based on Section 14.5 of *Numerical Recipes in C*, which is appended.

This is a "multimedia" introduction. These notes are not self-contained and complete; they are meant to be read along with Section 14.5 from *Numerical Recipes*, along with the accompanying material available at `www.linux.edu/physics/ph310/errors.html` This web page includes links to the datasets discussed in this document as well as links to Mathematica notebooks that implement the analysis of these data sets, along with some notes on using gnuplot for fitting.

II. SAMPLE DATA

A. Data Set 1

The first data set is a set of four measurements of the same physical quantity, say, for example, the gravitational field strength g. The measurements do not have the same uncertainties.

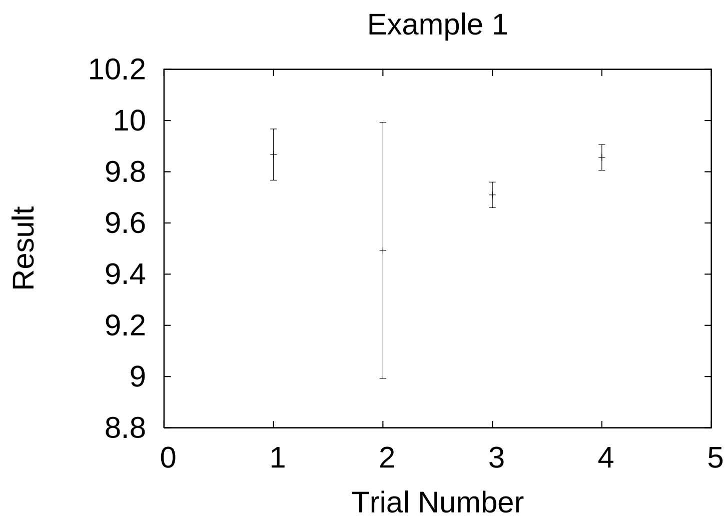

TABLE I: Sample data set 1

Trial #	g_i	σ_i
1	9.86689	0.1
2	9.49308	0.5
3	9.70960	0.05
4	9.85569	0.05

Questions:

1. Is is legitimate to combine the data to come up with a "best" value for g?

2. If it is, what is the "best" value of g given this data?

Comments: This is a well-known data analysis problem that is discussed in Section 7.2 of Taylor. Using the standard formulas given in Eqs. (7.10) and (7.11) we find that $g = 9.81 \pm 0.03\,\mathrm{m/s}^2$. Monte Carlo methods must, of course, agree with this.

B. Sample Data Set 2

Data set 2 is a standard data set in which measurements of the dependent variable y_i are made (with uncertainties σ_i) as the value of the independent variable x_i is varied. The uncertainties in the independent variables x_i are assumed to be negligible. Note that the uncertainties are not the same for all measurements. Data sets like this are often used to calibrate scientific apparatus. For example, you might be given the wavelengths of several known spectral lines, and use them to calibrate a spectrometer. In this case, the wavelengths are known to very high precision, so they correspond to the set of x_i's, and the y_i's might be pixel number on a CCD array in a spectrograph.

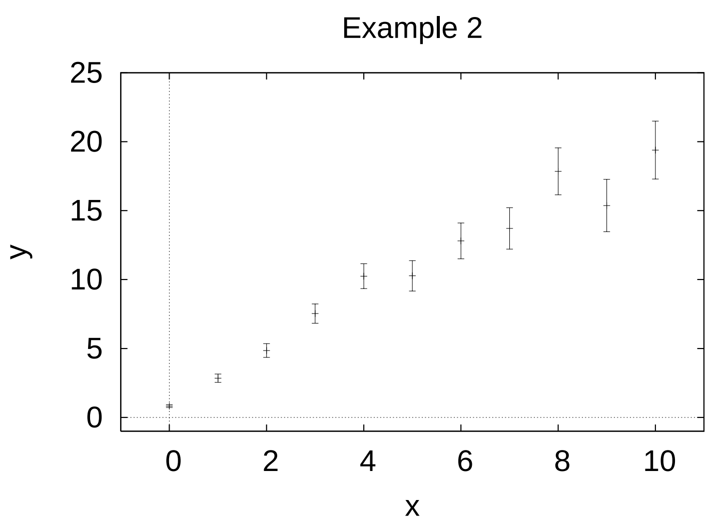

TABLE II: Sample data set 2

x_i	y_i	σ_i
0	0.8214	0.1
1	2.8471	0.3
2	4.8520	0.5
3	7.5347	0.7
4	10.2464	0.9
5	10.2707	1.1
6	12.8011	1.3
7	13.7108	1.5
8	17.8501	1.7
9	15.3667	1.9
10	19.3933	2.1

Questions:

1. Is there a linear relationship between x and y?

2. If there is, what are the best values of the slope and intercept?

3. If there is, what are the uncertainties in the slope and intercept?

4. Assume that this data set is to be used to calibrate a piece of apparatus, and you measure a new value of of the dependent variable Y (like the pixel number of an unknown spectral line in the spectrometer discussed above). What are the values and uncertainty in the value of the associated X (say, the wavelength of an unknown spectral line) that you determine from this measurement?

Comments:

- Analytical approaches to answering questions 2 and 3 are in every data analysis book. Formulas for the "best" slope and intercept are given in Eqs. (8.10-8.12) in Taylor, and the uncertainties are given in Eqs. (8.16) and (8.17), **BUT** these equations are only valid in the case that the uncertainties are the same for every data point. There are formulas for the case of unequally weighted points, but you won't find them in Taylor or other introductory books, although you could derive them pretty easily. A similar warning applies to the output from computer packages: unless you do something "special," the numbers you get are only strictly valid only for the case of data with uniform uncertainties.

- There is an analytical result for the answer to question 4, but I challenge you to find it in a book, especially for the case of unequal uncertainties. (This analytical result is more known by chemists than physicists it seems.) This uncertainty is definitely *not* given explicitly in the output of standard computer packages. The "difficulty" here is that the best fit slope and intercept are correlated, so naive addition of errors in quadrature won't work here.

  Notice that the line is "nailed down" pretty well on the lower left end of graph, but not so well at the upper right, and the uncertainty in X (given some measured Y) must reflect this. The uncertainty should also increase as the value of Y goes outside of the illustrated range.

C. Sample Data Set 3

The third data set is similar to the second, but this time the error bars on the individual points are smaller, and for convenience they are all equal in magnitude.

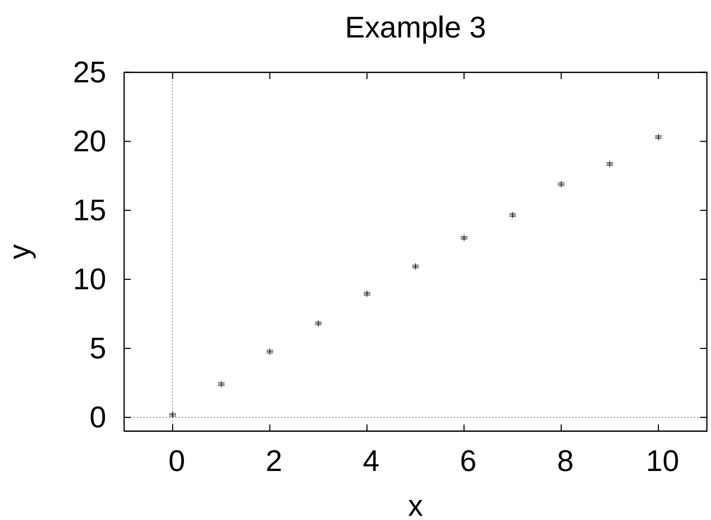

TABLE III: Sample data set 3

x_i	y_i	σ_i
0	0.188039	0.1
1	2.403512	0.1
2	4.76703	0.1
3	6.80721	0.1
4	8.95863	0.1
5	10.9312	0.1
6	13.0096	0.1
7	14.6587	0.1
8	16.8880	0.1
9	18.3569	0.1
10	20.3084	0.1

Questions:

1. What is the functional relationship between x and y? Is it linear? If it is, what are the best values of the slope and the intercept?

2. Assume that this data set is to be used to calibrate a piece of apparatus, and you measure a new value of y, (like the pixel number of an unknown spectral line in the spectrometer discusses above). What is the best value of x that you determine from this measurement, and what is its uncertainty?

III. "UNIVERSAL" PROCEDURE

The procedure here follows that described in the appended Section 14.5 of *Numerical Recipes*. As above, assume that we are given N data points, (x_i, y_i), and associated uncertainties σ_i (and assume that uncertainties in the values x_i is negligible). We fit the data to some model function $f(x;\mathbf{a})$, where the $\mathbf{a}$ is a "vector" containing the adjustable parameters of the model. If the model is good, we assume that that there is some true (but unknown) set of parameters $\mathbf{a}_{\text{true}}$.

1. Assume a trial functional relationship between x and y:

$$y = f(x; a_1, a_2, a_3, \dots) = f(x; \mathbf{a})$$

   For sample data set 1 the trial relationship would be

$$y = a_1 = \text{constant};$$

   For sample data set 2 the trial relationship would be linear:

$$y = a_1 + a_2 x.$$

2. Pragmatic advice: Plot function $y = f(x;\mathbf{a})$ along with the data on the same graph and adjust the parameters of $\mathbf{a}$ to get an approximate fit.

3. Find the values a_1, a_2, ... that minimize χ^2, where

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - f(x_i;\mathbf{a})}{\sigma_i} \right)^2 .$$

   This is what software packages can do very efficiently for you. Let's call this set of best fit parameters $\mathbf{a}_0$.

4. Replot $y = f(x;\mathbf{a}_0)$ and data on same graph — does the fit "look" good?

5. If the fit looks good, plot the residuals, with error bars on the data points, i.e., plot $(y_i - f(x_i;\mathbf{a}_0))$ vs. x_i. Do the residuals look randomly distributed about zero? Are about 2/3 of the points within one error bar of zero?

6. If the fit looks good, and the residuals look randomly distributed, find the value of the reduced χ^2, or χ^2_R. Is it close to 1?

7. If you now have confidence in your model, proceed to estimate the errors in your fit parameters using the following steps; if not, think hard about other ways to estimate uncertainties.

8. Generate a hypothetical data set using your set of fit parameters $\mathbf{a}_0$ as a reasonable approximation to $\mathbf{a}_{\text{true}}$, along with randomly selected "noise" that is consistent with what you believe to be true about your errors. This is easy to do using modern software packages. (See the *Mathematica* notebook `statistics_tools.nb`) Calculate a new set of best fit parameters $\mathbf{a}_1$ for this hypothetical data set. (If you're using this set of parameters to calculate something else, as in the case of calibrating an instrument for the determination of an unknown value, go ahead and perform the calculation using the parameters $\mathbf{a}_1$.)

9. Repeat the previous step many times, generating parameter sets $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a_4}$, etc. The accumulated results from the hypothetical data sets won't give you any better sense of the value of $\mathbf{a}_{\text{true}}$, but they will tell you about the spread around $\mathbf{a}_{\text{true}}$ that arises from the uncertainties in your measurements (assuming that $\mathbf{a}_0 \simeq \mathbf{a}_{\text{true}}$).

10. Calculate the standard deviation of the parameter of interest from your accumulated data from hypothetical data sets; this is your uncertainty.

IV. MATHEMATICA EXAMPLES

- `statistics_tools.nb` — An introduction to some of the packages that you need to load, and some of the tools that you need to use to perform statistical analysis.

- `weighted_mean.nb` — An analysis of the sample data set 1.

- `line_fit.nb` — An analysis of the sample data set 2. This notebook also includes examples of how to generate contour plots of $\Delta\chi^2$ like those illustrated in Fig. 14.5.3 of *Numerical Recipes.* It also includes a section illustrating how to get the uncertainty using the covariance matrix, as well as an illustration of how a naive approach to the calibration problem can go wrong.

- `quadratic_fit.nb` — An analysis of sample data set 3, focused on using the data as a calibration data for a scientific instrument.