PHYS-3901 Intermediate Lab
Notes on the Treatment of Errors

J.W. Martin
The University of Winnipeg

October 26, 2011

Abstract

I review the treatment of statistical and systematic errors, fitting, and uncorrelated and correlated errors.

1 Statistical Error in Counting Experiments

Imagine a counting experiment where data is acquired for a fixed time t. Suppose we know the average rate in the experiment should be about R. The probability of measuring N counts in the experiment is given by the Poisson distribution:

$$f(N;Rt) = \frac{(Rt)^N e^{-Rt}}{N!}. \tag{1}$$

Therefore, repeating this experiment an infinite number of times, we would measure, on average $\langle N \rangle = Rt$ counts. Histogramming the results of those experiments we would build up a Poisson distribution, centered on $\langle N \rangle$, with a width given by the standard deviation $\sigma = \sqrt{\langle N \rangle}$.

What we take this to mean in our class is: if we measure a certain number of counts in our experiment N, the statistical error in the number of counts is:

$$\delta N = \sqrt{N}. \tag{2}$$

2 Fitting with Uncorrelated Errors

Suppose we desire to fit data in the form x_i, y_i, σ_i, where σ_i is the uncorrelated error in the measurement y_i. Suppose the fit function is linear $f(x) = mx + b$, and by fitting we desire to determine the parameters m and b, and their errors.

We first construct the variable χ^2 to estimate the goodness of fit [2,4]:

$$\chi^2 = \sum \frac{(y_i - f(x_i))^2}{\sigma_i^2}, \tag{3}$$

where the sum is over the data points.

The procedure of least-squares fitting is to minimize χ^2. In the case of fitting a line, we'd take derivatives of χ^2 with respect to m and b and set them equal to zero, and then solve for the best m and b. Doing this gives:

$$b = \frac{1}{\Delta}\left(\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2}\right) \tag{4}$$

$$m = \frac{1}{\Delta}\left(\sum \frac{1}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2}\right) \tag{5}$$

where

$$\Delta = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2}\right)^2. \tag{6}$$

Note that these reduce to the equations you used in 2nd year [1], if all the errors σ_i are equal.

On top of this, by evaluating the second derivative of χ^2, we can estimate the uncorrelated errors in the intercept and slope, respectively:

$$\sigma_b^2 = \frac{1}{\Delta} \sum \frac{x_i^2}{\sigma_i^2} \tag{7}$$

$$\sigma_m^2 = \frac{1}{\Delta} \sum \frac{1}{\sigma_i^2} \tag{8}$$

Finally, there are programs out there that can fit arbitrary functions $f(x)$ to your data, with any number of parameters. The programs I normally use are available for free from CERN [5, 6]. Common functions you might want to fit are polynomials, exponentials, or Gaussians, or any sum or combination of them. The way these programs work are to search the parameter space to find the minimum χ^2. This gives the "best fit" for the

parameters. In order to determine errors, they then evaluate the second derivative of χ^2 numerically, by varying the parameters slightly around the minimum.

It is therefore fine with me if you would prefer instead of using the equations above to use one of these programs, even for linear fits.

3 Goodness of Fit

The minimum value of χ^2 found by the minimization process can be interpreted in terms of a goodness of fit.

For this, it is useful to define the number of degrees of freedom of the fit:

$$\nu = (\text{number of data points}) - (\text{number of fit parameters}), \tag{9}$$

which in turn enables definition of the *reduced* χ^2:

$$\chi^2_\nu = \frac{\chi^2}{\nu}. \tag{10}$$

The value of the reduced chi-square χ^2_ν relates to the goodness of fit:

- If the value is $\chi^2_\nu \sim 1$, it means the fit is "good" or "reasonable".
- If χ^2_ν is much larger than unity, it means the quality of the fit is "bad" or "poor", or that the function being fit is not an accurate representation of the data.
- If χ^2_ν is much smaller than unity, it means the quality of the fit is "too good", or that the errors have been overestimated. Or that the data appear to agree suspiciously too well with the function begin fitted.

With enough experience, the value of χ^2_ν and goodness of fit can be estimated graphically from the data. If the best fit line, when plotted on top of the data with error bars, passes perfectly through each point and well within the error bars, then clearly χ^2_ν will be less than 1 and the fit is "too good". If the data points never touch the best fit line within the error bar, then clearly χ^2_ν will be greater than 1 and the fit is "bad". If, in a fit to say 10 points, a few of the data point have error bars that don't touch the best-fit line, but the rest of them do touch the best fit line (some of them possibly lying perfectly on the best fit line so that their contribution to χ^2_ν would be close to zero) then the fit is "good" and you can anticipate $\chi^2_\nu \sim 1$.

Finally, there are varying degrees of "too good", "good", and "bad", and some people (not me) prefer to state "the probability of exceeding χ^2" which can be looked up in tables (see e.g. Appendix C of Ref. [2]). One way I like to interpret the meaning of this probability is: if I took these data points and their errors at face value, and then randomly moved the data points around within their error bars, what's the likelihood I would get something close to this value of χ^2 again? Ideally, this probability would be around 50%.

The interesting thing about the probabilities is that, the larger the number of degrees of freedom, the sharper this distribution becomes. So if you have 10 degrees of freedom and you measure $\chi^2_\nu = 1.05$, the probability is about 40% (a slightly poor fit). But if you have 200 degrees of freedom and you measure the same $\chi^2_\nu = 1.05$, it turns out the probability is smaller 30%. It's actually not that unlikely to get $\chi^2_\nu = 1.5$ for 10 data points (prob $\sim$ 15%), but for 200 data points, this would be a disaster (prob $< 0.1\%$).

So the rule of $\chi^2_\nu \sim 1$ is a little more qualitative than looking up the probability in the table.

4 Correlated Errors

Imagine you were taking data, and you made a random error each time. This error would be uncorrelated to the previous measurement you made. You could then assign an appropriate error and be confident that the equations in the previous sections would work. This is indeed the case for the $\sqrt{N}$ error relating to counting experiments.

However, imagine instead that you made the *same* error every single time you made the measurement.

For example, say you measured the length of a number of lines using a ruler, and you knew the lengths should be multiples of one another, and got values 1 cm, 2 cm, 3 cm, etc. You then decide you fit your data to the function $f(x) = mx + b$ where x_i is the suspected number of multiples and y_i is the length you measured. You arrive at the values $m = 1$ cm and $b = 0$. And you get errors σ_m and σ_b dependent on what vertical error bars σ_i you felt were reasonable based on how well you could read the ruler. You also get $\chi^2_\nu = 1.0001$. The fit is "good".

However, suppose you made one error consistently throughout this whole process. You misread the scale on the ruler. It was actually in inches! Therefore you made a factor of 2.54 error in every point you measured.

Should you go back and increase the error on every single measurement you made to some huge value? If you did that, χ^2_ν would become very small

This is an example of an error that is *correlated* between the data points. It is not random or *uncorrelated*. The best way (for us) to deal with these kinds of errors is to treat them separately from other possible errors.

In this case, we would probably go back and correct each data point for the error we made (changing cm to inches). If there was some uncertainty associated with doing this, it would probably not be assigned as an uncorrelated error to be inluded in the σ_i used in the fitting process.

5 Systematic Errors

The above example also relates to a form of a systematic error. You made a mistake about something systematically relating to each points.

Systematic errors, however, can also be uncorrelated.

Imagine, in the previous example, that after taking all the measurements, you also realized you weren't that careful about lining up the zero point on your ruler, and that you didn't take this into account when making the measurements, or assigning the error to the measurements. This would at first manifest itself as a poor χ^2_ν, giving you a hint that you made some error.

After realizing this, here's a strategy you could take to address this systematic error. Measure just one of the lines using whatever technique you were using before. Then measure more carefully, specifically addressing the systematic error you made. Take the difference between the two measurements as the likely random error you made for each of the lines. This error could then be added *in quadrature* to the error you previously assigned, giving a revised random error for every data point. You could then perform your fit again, and χ^2_ν would hopefully improve.

This extended example exemplifies two things about experimental physics, one positive and one negative.

- The positive: The best way to address systematic errors is to change something about the experiment you did and then investigate carefully what happens in the experimental result. Consider carefully whether you would have made the same error (correlated) or a different error (uncorrelated) on each of the data points you measured.
- The negative: There can be a psychological effect in experimental physics where, as soon as $\chi^2_\nu \sim 1$, you stop looking for uncorrelated errors. The example I gave above is

an example of this. Doing a study in the order given above can lead you dangerously down this path, especially if you take χ^2_ν as some measure of success. This might, for example, lead you to overestimate your errors in one area, while completely neglecting the real error at hand.

One way to avoid the negative point is first to make a list of all the systematic errors you think you have made. Then do experiments to limit those errors, never considering χ^2_ν. Then, after all is said and done, analyze the data and determine χ^2_ν. Basically, be aware of potential errors you could be making, learn about them, but do not necessarily use χ^2_ν as a real measure of success.

6 Propagation of Errors

The formulae used in 2nd year [1] were correct, as long as they are used for uncorrelated errors. For example, suppose you measure the physical quantities a and b and assign uncorrelated errors δa and δb. You then desire to calculate the value of some function $f(a, b)$. The error in $f(a, b)$ is:

$$\delta f^2 = \left(\frac{\partial f}{\partial a}\right)^2 \delta a^2 + \left(\frac{\partial f}{\partial b}\right)^2 \delta b^2. \tag{11}$$

For correlated errors it is more complicated. Consider the case above where a and b are exactly the same physical quantity a. And the function is $f(a) = a$. You cannot magically reduce your errors by a factor of $\sqrt{2}$, as the above formula would imply! The correct error is obviously δa.

In the case of correlated errors, the correct way to do things is to consider carefully the correlations between the measured quantities and their errors. If I change a, does b automatically change in some way? If the answer is yes, then clearly I cannot vary $\frac{\partial f}{\partial b}$ without necessarily varying $\frac{\partial f}{\partial a}$. In this case, the error in a is manifested also as an error in b and they are not really independent. Their errors will therefore also not likely be uncorrelated.

The answer "sort of" is also possible; b can be somewhat correlated with a. In such cases, a full consideration of cross terms containing e.g. $\frac{\partial^2 f}{\partial a \partial b}$ would be necessary. We will not attempt such considerations in this class. Either it is fully correlated (b is a well-defined function of a), or it is completed uncorrelated. Such strategies are also best followed in real life, too.

7 References

[1] Your 2nd-year Lab manual.

[2] P. R. Bevington, *Data Reduction and Error Analysis in the Physical Sciences.*

[3] W. R. Leo

[4] Taylor

[5] http://root.cern.ch

[6] http://wwwasd.web.cern.ch/wwwasd/paw/