Random Uncertainties

1 Introduction

All measurements come with a degree of uncertainty that is sometimes referred to as *error*. Error, in this case, does not indicate a mistake, rather it is a concept used to quantify the confidence that can be placed in a measurement.

Uncertainties are often subdivided into two categories: *random* and *systematic*. Random errors can be reduced by repeating a measurement several times. Systematic errors cannot. Another type of systematic error is *instrument error* or the error associated with the tool that is used to make a measurement. This is sometimes referred to as the instrument *resolution*.

In this lab we will explore random errors and statistics in a counting experiment.

2 Learning Objectives

At the conclusion of this activity you should be able to:

- Understand the concept of random error.
- Be able to use statistical analysis tools like average and standard deviation.
- Experimentally verify that the uncertainty in a counting experiment is given by the square root of the number of counts.

3 Background

3.1 Reading Assignment

Chapters 11.1-11.2 (pp. 245-252) in Taylor, J. R. (1997) *An Introduction to Error Analysis*. Sausalito, CA: University Science Books.

3.2 Statistics Primer

Statistical analysis is a large and detailed sub-field of error analysis that we cannot hope to adequately treat in an introductory laboratory course. The purpose of this section is to give you a brief introduction to the concepts that you will need to apply the ideas of basic statistical analysis to your measurements.

There are several excellent references to consult for a more coherent and detailed introduction to error analysis. A few are listed below;

- John R. Taylor, *An Introduction to Error Analysis*, 2nd ed. (University Science Books): Chapter 4 - "Statistical Analysis of Random Uncertainties" is a great place to start. See also Chapter 11 – "The Poisson Distribution".
- Clifford E. Swartz, *Used Math*, 2nd ed. (AAPT): Chapter 5 - "Statistics" is a more rigorous treatment of the mathematics behind statistical analysis.

The following is a brief discussion of key concepts and definitions.

Sample vs. Population: In the study of statistics, the words *population* and *sample* have technical definitions that are perhaps different from their common usage in English. In statistics, the *population* refers to the complete set of things that are being studied. The *sample* on the other hand is a set of data that have been collected or selected from a population using a set of selection criteria.

For example, if we want to measure the height of students in our lab class, the *population* could be all human beings. The *sample* would be the subset of all human beings who happen to be registered for this lab section.

The distinction between a population and a sample is subtle, but important. Often, the criteria that is used to select a sample will introduce a bias. For example, if we attempt to measure the average height of a human being based on a sample of NBA basketball players or measure the life span of a human being based on a sample of people who currently live in retirement homes. This type of bias is often referred to as *selection bias* and can be a source of systematic error in a measurement.

Mean or Average: The mean or average of a sample is often the best estimate we can quote for a given quantity. The mean is calculated using the familiar formula;

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_N}{N} = \frac{\sum x_i}{N} \tag{3.1}$$

where N is the number of individual measurements, x_i.

In most spreadsheet programs the mean can be calculated in a cell using the `=AVERAGE(sample)` function. Here, `sample`, is the cell range containing the values whose average we wish to calculate.

Standard Deviation: The standard deviation is an estimate of the average uncertainty of the separate measurements used to calculate the mean of a sample. The standard deviation σ_x of the sample is given by:

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum \left(x_i - \bar{x}\right)^2} \tag{3.2}$$

The standard deviation can be calculated in a spreadsheet by using the `=STDEV(sample)` function.

3.3 Normal Distribution

Also known as a *bell curve* or *Gaussian distribution*, the normal distribution is a probability function that is used to model the shape – or uncertainty – associated with a given data sample. The normal distribution is given by:

$$G_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\bar{x})^2/2\sigma_x^2} \tag{3.3}$$

where $G_\sigma(x)$ gives the probability of a data point being found at position x, $\bar{x}$ is the central value (or mean), and σ is the width (or standard deviation) of the distribution. Equation 3.3 is plotted in Figure 3.1.

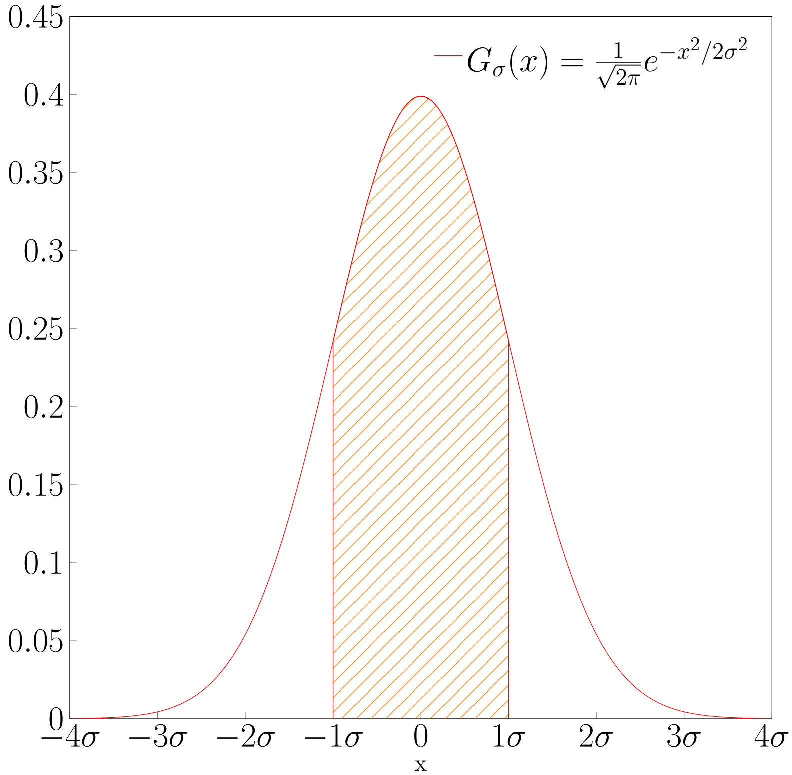

Figure 3.1: The function, $G_\sigma(x)$, represents a Gaussian distribution centered on zero ($\bar{x} = 0$) with a width of σ. The shaded area, between $-\sigma$ and $+\sigma$, is the probability that a single measurement will fall within one standard deviation of the mean. For a Gaussian distribution, this area corresponds to about 68% of the total area bounded by the function.

For data that are accurately modeled by a Gaussian distribution, we expect that 68% of the events will be bounded by the domain between $-\sigma$ and $+\sigma$. Table 3.1 lists the percent of events that will fall within n standard deviations of the mean.

$\pm n\sigma$	1	2	3	4
probability (%)	68.3	95.4	99.7	99.99

Table 3.1: The probability that a single measurement will fall within n standard deviations of the central value.

See Taylor, Chapter 5.3 for more information.

3.4 COUNT RATE

A count rate R is the number of events ν counted in a given time interval T:

$$R = \frac{\nu}{T}. \tag{3.4}$$

If the count rate is measured repeatedly, we may want to compute the average of those measurements. We define μ to be the average number of observed events.

$$\mu \equiv \bar{\nu}. \tag{3.5}$$

The average count rate is then given by:

$$\bar{R} = \frac{\mu}{T}. \tag{3.6}$$

3.5 UNCERTAINTY IN A COUNTING EXPERIMENT

In a counting experiment an observer records the number of events that occur in a given time interval as described in Section 3.4.

For example:is given by:

$$\sigma_\nu = \sqrt{\mu}. \tag{3.7}$$

4 Radiation Safety

The Cesium-137 sources used in this project are weak and pose minimal health risk. However, as a matter of safety, one should observe standard precautions when handling these (and all other) radioactive materials.

- Avoid unnecessary handling. Do not break the plastic encapsulation.
- Do not eat, drink, or smoke near the radioactive materials.
- Radioactive materials should be stored in the clearly labeled, lead container when not in use.
- Immediately report any loss or breakage to your instructor, TA, or laboratory manager.

Nuclear radioactive decay produces three types of radiation: α particles (helium nuclei), β particles (high energy electrons and positrons), and γ rays (high energy photons). The Cesium sources that are used in this lab emit only β and γ radiation.

5 Procedure

The goal of this lab is to experimentally investigate the relationship given in Equation 3.7:

$$\sigma_\nu = \sqrt{\mu}.$$

A radioactive source is used to generate data. Radioactive decays occur with a definite average lifetime but the decay of any particular nucleus is random and does not depend on the decay of any other nucleus. These random events are a perfect place to study statistics and uncertainty distributions in a counting experiment.

The radioactive decays will be measured using a Vernier Digital Radiation Monitor* which consists of a Geiger-Muller tube and a digital counter.

This experiment requires the use of the Vernier Logger Pro software package. Instructions on how to download and install the program are available on Blackboard: `Lab Assignments` ▸ `Reference Material` ▸ `Vernier Logger Pro Installation Instructions`.

5.1 Data Collection

For a fixed rate R_o, the time interval T_i will be varied to obtain different average counts μ_i.

- The average number of counts μ_i will be recorded for each experiment with time interval T_i.
- The standard deviation σ_{ν_i} will be calculated directly from the data of each experiment.
- Finally we will plot these data to test the relationship given in Equation 3.7.

The following steps are intended to direct your efforts as you develop and create your own experiment.

1. Open the Logger Pro program on your computer.
2. Connect the LabPro Data Acquisition unit (DAQ) to your computer with the USB cable.

*See http://www.vernier.com/products/sensors/radiation-monitors/drm-btd/

3. Connect the Digital Radiation Monitor to LabPro using one of the two Digi/Sonic ports. Turn on the radiation monitor. The timer switch should be set to the “CPM/CPS” setting.

4. In Logger Pro, use Experiment ⟩ Set Up Sensors to confirm the the radiation monitor is recognized and is assigned to the correct Digi/Sonic port.

5. In Logger Pro, select the following file: `File ▸ Open... ▸ Nuclear Radiation w ▸ Vernier ▸ 04 Statistics.cmbl`.

6. Position the Cs-137 sample in front of the digital radiation monitor such that you measure an average count rate† of about 100 counts per second:

$$\overline{R_o} \approx 100 \text{ counts/sec.} = \frac{100 \text{ counts}}{1 \text{ sec.}}$$

   Leave the Cs-137 sample at this same location for the rest of the experiment.

7. Use Analyze ⟩ Statistics to calculate the observed average count rate $\overline{R_o}$ and standard deviation σ_v of the data.

8. Using the measured $\overline{R_o}$ and Equation 3.5, predict the time T required such that the average number of counts μ will be equal to 1.

9. Adjust the “sampling rate” in Logger Pro using the value of T calculated above. Use Experiment ⟩ Data Collection... (also available via the stopwatch icon) to adjust the “Sampling Rate”. Note that the T represents the *sampling rate* and can be expressed in samples/second, Hertz (Hz), or seconds per sample.

10. Take data for 30 seconds to confirm that the observed average number of counts μ is approximately 1 during the time interval. Is the observed count rate $\mu \pm \sigma_v$ compatible with 1 count per time interval? Note that μ and σ_v are both automatically calculated and reported in the Analyze ⟩ Statistics box.

11. Calculate several different “sampling rates” required to measure distributions with predicted average number of counts, μ between 1 and 30 counts per interval. How many different “sampling rates” will you use to measure the space between 1 and 30 counts per interval?

12. Take data at each of your proposed points. For each trial, record the measured average counts per time interval μ and the standard deviation σ_v.

†Vernier digital sensors are limited to a maximum data rate of 250 Hz. The Digital Radiation Monitor is recommended for use below 100 Hz.

5.2 Analysis Suggestions

- Plot the data such that there exists a *predicted* linear relationship between μ and σ_v (see Equation 3.7). Remember to include error bars.
- Fit the data with a linear model to quantify the prediction.
- Do your data support the relationship suggested in Equation 3.7?

6 Lab Notebook

Your submission will be evaluated using the following rubric:

Lab Notebook Practices

- Lab Notebook Mechanics (4 points)
  - Relevant information *e.g.*: your name, your lab partner's name, date, *etc.* is present.
  - The notebook is organized and easy to read. Markdown cells are used for narrative text. Code cells are clearly organized and commented.
  - The ZIP file of the notebook is healthy and runs correctly.
- Experiment Purpose (2 points)
  - In your own words, state the purpose of this experiment.
  - The work you record in your notebook should specifically address the stated purpose.
- Calculations & Error Propagation (6 points)
  - Present the equations used to analyze your collected data.
  - Present the equations used to account for and propagate error.
- Data Collection & Analysis (8 points)
  - The notebook tells a scientific story; it is an accurate record of the work that you did.
  - Record the informal observations you made during the experiment.
  - The notebook should show evidence of trial and error. Keep a good record of your work – recording mistakes is useful.
  - *Briefly* describe the methods you used to collect your data.
  - Measured quantities – that do not appear directly in your plot(s) – are clearly recorded.
  - Record rough data and plots that you used to verify that the analysis was on the right track.

Results and Interpretation

- Results (5 points)
  - Clearly state the final result(s) of your experiment. Remember to quote your result with units and appropriate significant digits.
  - Final result plots are well formatted and meet the standards described in the Figure Formatting reference.
- Physical Interpretation (2 points)
  - Interpret your result in terms of the underlying physics.
- Significance (5 points)
  - Compare measured results to each other and/or to a known/expected value.
  - Choose the best available tools for your comparison (*e.g.* plots, pictures, discrepancy, significance, etc).
- Confidence (8 points)
  - Communicate to your audience how seriously your result should be taken. How confident you are in your result?
  - Discuss factors that may be affecting the *accuracy* and *precision* of your result.
  - Suggest improvements to your experiment to address your confidence, accuracy, and precision.