University of Rochester

Laboratory I
Understanding Statistical Uncertainties

DEPARTMENT OF PHYSICS & ASTRONOMY

PHYSICS 113 - 121 - 181

GENERAL PHYSICS I AND MECHANICS

Name: ______________________________ **Date:** ______________________

Collaborators: ____________________________ **Lab Section:** ____________

PRELAB EXERCISES (6 points)

This prelab must be completed and handed in to the lab TA at the start of the lab.

Question 1 **2 points**

You are trying to see if there is a pattern to when spam arrives in your mailbox, so you count the number of spam emails received over the last month and bin them by time of day. Out of 1595 spam emails, 784 were received during the day, between 6 AM and 6 PM. Treating spam emails as Poisson distributed, do your data support the idea that spammers prefer one part of the day over the other? *Hint: add error bars to the data and see if there is overlap.*

Question 2 **2 points**

You investigate further and subdivide the day into four equal time intervals. What conclusion do you draw from your data now, and why?

Time of day	No. of spam emails
midnight to 6:00 AM	383
6:01 AM to noon	401
noon to 6:00 PM	374
6:01 PM to midnight	447

Question 3 **2 points**

You are a quality engineer for a paper company and must ensure that the area of A4 sheets being produced has less than 0.1% error. You analyze a sample of A4 paper and find the width and length to be, respectively, 210 ± 0.2 mm and 297 ± 0.3 mm. Does the company have a problem? Explain.

Objective

Develop familiarity with the statistical tools used to describe experimental data, including mean, standard deviation, and error estimators. Apply these tools to estimate measurement uncertainties and propagate them through derived quantities such as areas and ratios.

Theory

Introduction

Throughout this semester you will perform experiments that involve measurements. No measuring device is perfect, so none will give you infinitely precise results. Every measurement contains the true value you seek, obscured by errors introduced by these imperfections. Understanding errors is therefore crucial for extracting the true value from imperfect data. Some errors are large, some small — understanding them means understanding how they are distributed.

You are probably familiar with elections and polling. During an election, each voter has a preference; there is a *distribution* of preferences across the population. The election reveals that distribution, and a *statistic* — the winner — is determined. Prior to an election, polls attempt to predict this statistic: they are *estimators* of it. They do so by *sampling* a fraction of the population. If the sample is chosen to represent the population faithfully, the estimator is *unbiased*. This does not mean the poll predicts the election result exactly, but rather that the error in the prediction depends on the sample size. The larger the sample, the closer the prediction is to the actual result, and the smaller the *statistical error*. If the sample is not chosen properly, the estimator is *biased*. The prediction will then carry a *systematic error* that does not shrink as the sample size grows.

The difference between a good and a bad experiment comes down to understanding systematic and statistical errors. Distributions are useful tools in experiments that involve counting. When counting *discrete* quantities — such as the number of lightning bolts per hour during a storm — a *probability distribution* gives the chance of observing one, two, or even a hundred strikes. These are typically written as p_n, where n runs over the discrete possibilities. When dealing with *continuous* quantities — such as the time between successive lightning strikes — a *probability density function* (pdf) gives the chance of an event occurring within any time interval, such as within the first eight minutes, after more than 10.5 minutes, or somewhere between the two. These are typically written as $\rho(x)$, where x runs over the continuous values.

Common Statistics and their Unbiased Estimators

The most common and useful statistic is the *mean* (or *average*). Let X be the collection of all possible outcomes in an experiment. A particular outcome is written $x \in X$. The mean for discrete or continuous distributions is, respectively,

$$\mu = \sum_{x \in X} p_x x \quad \text{or} \quad \int_X dx\, \rho(x) x.$$

The spread of a distribution is captured by the second most important statistic, the *variance*,

$$\sigma^2 = \sum_{x \in X} p_x (x-\mu)^2 \quad \text{or} \quad \int_X dx\, \rho(x)(x-\mu)^2,$$

the square root of which, σ, is the *standard deviation*. In other contexts you will see these statistics written using *expectation* notation, $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathbb{E}[(X-\mathbb{E}[X])^2]$, or with angle brackets, $\mu = \langle x \rangle$ and $\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle$. When there are multiple types of outcomes, say X and Y, subscripts distinguish the means, μ_X and μ_Y, and variances, σ_X^2 and σ_Y^2.

When X and Y are dependent random variables, our final important statistic is the *covariance*,

$$\sigma_{XY} = \sum_{x \in X} \sum_{y \in Y} p_{x,y}(x-\mu_X)(y-\mu_Y) \quad \text{or} \quad \int_X dx \int_Y dy\, \rho(x,y)(x-\mu_X)(y-\mu_Y).$$

A positive covariance means that when X is above its mean, Y tends to be above its mean as well, while a negative covariance indicates the opposite. When X and Y are independent, $\sigma_{XY} = 0$.

The formulas above are useful for theoretical purposes, but not directly for experimental ones. Probability distributions and pdfs — like the true distribution of voter preferences before an election — are generally not known in advance, so you usually cannot compute these statistics directly. Fortunately, a sample of N measurements, $s = \{x_1, x_2, \ldots, x_N\}$, can be used to estimate them. We call the collection s *data*. The unbiased estimator of the mean is

$$\hat{\mu} = \frac{1}{N} \sum_{n=1}^{N} x_n, \tag{1}$$

and for the variance,

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{n=1}^{N} (x_n - \hat{\mu})^2. \tag{2}$$

The little hats (circumflexes) distinguish estimators from the true statistics. The estimated standard deviation is typically quoted alongside the estimated mean as $\hat{\mu} \pm \hat{\sigma}$, $\hat{\mu}(\pm\hat{\sigma})$, or $\hat{\mu}^{\pm\hat{\sigma}}$. These two quantities are what you will compute from your data. Make sure you are comfortable with the formulas before the lab begins — you will use them throughout the course.

Error Propagation

Suppose you measure x and y, so your data consist of pairs (x_i, y_i) for $i = 1, 2, \ldots, N$. You are interested in a derived quantity z that depends on x and y through some function f,

$$z = f(x, y). \tag{3}$$

Since the quantities are related, their statistics are related as well. Given estimates $\hat{\mu}_x \pm \hat{\sigma}_x$ and $\hat{\mu}_y \pm \hat{\sigma}_y$, can you obtain $\hat{\mu}_z \pm \hat{\sigma}_z$?

If the estimated standard deviations of x and y are both small compared to their means, $\hat{\sigma}/\hat{\mu} < 0.1$, and f is reasonably well-behaved, then yes. The mean of z is estimated as $\hat{\mu}_z = f(\hat{\mu}_x, \hat{\mu}_y)$. The variance is

$$\hat{\sigma}_z^2 \approx \left(\frac{\partial f}{\partial x}\right)^2 \hat{\sigma}_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \hat{\sigma}_y^2, \tag{4}$$

with the partial derivatives evaluated at the estimated means.

Special Distributions

Poisson Distribution

The Poisson distribution is one of the most important discrete distributions. It applies when counting events that occur independently at a fixed average rate. Consider a lightning storm in which it is reasonable to treat each strike as independent. Over the full storm, strikes occur at some average rate α. If we divide the storm into time intervals of length Δt, we expect on average $N = \alpha \Delta t$ strikes per interval. Since the strikes are random and independent, some intervals will contain more than N strikes and others fewer. The probability of observing exactly n strikes in an interval is

$$p_n = \frac{N^n}{n!} e^{-N}. \tag{5}$$

Since both the mean and variance of this distribution equal N, you expect $N \pm \sqrt{N}$ strikes per interval. This is a useful rule of thumb: in any counting experiment, a quick estimate of the statistical error is simply the square root of the count.

Gaussian Distribution

The Gaussian distribution is also called the normal distribution or bell curve. The probability density of a Gaussian random variable is

$$\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{6}$$

where μ is the mean and σ is the standard deviation. The distribution plays an important role when combining results from multiple experiments. In each experiment you collect data and estimate the mean $\bar{x}$. If you repeat the experiment N times and obtain N such sample means, how are those means distributed? According to the central limit theorem, they follow a Gaussian distribution with variance $\sigma_\mu^2 = \sigma^2/N$.

Equipment

- 1/2 meter stick
- Geiger Counter & Wand

Experiment

Experiment 1: The Psychophysics of Reacting to Visual Stimuli

In this experiment you and your partner will measure each other's reaction times without a stopwatch. One partner drops a ruler while the other catches it. The position of the catching hand records how far the ruler fell. Since the ruler starts from rest, the distance h it falls in time t obeys the kinematic formula

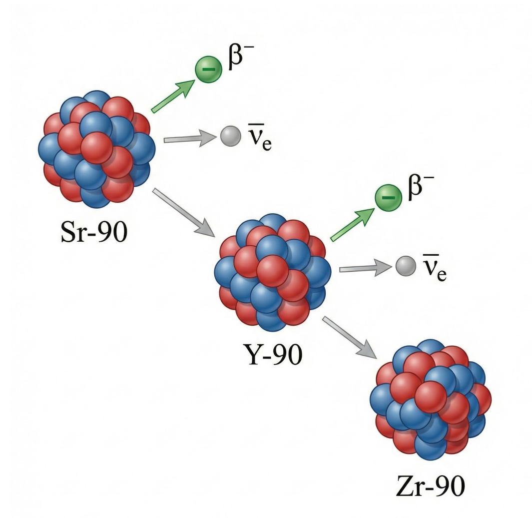
$$h = \frac{1}{2} g t^2 \quad \text{where} \quad g = 9.81\,\frac{\text{m}}{\text{s}^2} \tag{7}$$

is the gravitational acceleration at the surface of the Earth. A measurement of h therefore yields a measurement of t.

1. You and your partner choose the unique roles of either **Lab Partner A** or **Lab Partner B**.
2. **Lab Partner A**: Hold the ruler vertically with the 50 cm end at the top, as high up as possible.
3. **Lab Partner B**: Position your open hand at the lower end of the ruler, not touching it. Line up the top of your hand with the 0 cm mark, say "Ready," and begin waiting. Keep your eyes on your hand.
4. **Lab Partner A**: After hearing your partner say "Ready," respond with "Set" and wait however long you like. Without warning, release the ruler.
5. **Lab Partner B**: As soon as you see the ruler begin to fall, close your hand to catch it. Once caught, show your clenched hand to your partner. (Practice steps 3–5 until you are both comfortable with the procedure.)
6. **Lab Partner A**: Find the ruler mark just visible at the top of your partner's clenched hand and record it as h_1 in the post-lab.
7. Repeat steps 2–6 twenty-four more times, recording $h_2, h_3, \ldots, h_{25}$.
8. Switch roles and repeat steps 2–7.
9. Complete Experiment 2 before working on the post-lab questions.

Experiment 2: Radioactive β-Decay

In this experiment you will count the number of β-particles (electrons) emitted by a radioactively decaying ^{90}Sr (strontium-90) source. ^{90}Sr has a half-life of 28.8 years and decays into ^{90}Y (yttrium-90) by emitting an electron (β^-) and an antineutrino ($\overline{\nu}_e$). ^{90}Y has a half-life of 64 hours and decays into stable ^{90}Zr, emitting another β^- and $\overline{\nu}_e$. It is these β-particles that you will be counting.



The ^{90}Sr sample is a small button sealed within a protective steel frame. The frame blocks the vast majority of emitted electrons, so your net radiation exposure will be less than that of a cross-country flight. A Geiger counter with an external wand measures the radiation leaving the sample. Since the sample emits β particles at a fixed rate and each emission is independent of the last, the process is well described by a Poisson distribution. You will record counts per second in three different shielding configurations and examine how the statistical uncertainty depends on the number of trials.

1. Turn the Geiger counter on. It will make a rapid clicking sound when placed near the sample; you may mute it if you wish. Set it to take 1-second readings and test it by placing the wand close to the source with nothing in between. If you read fewer than 100 counts per second, consult your TA.
2. Examine the steel box containing the sample and try inserting the different plates and the wand. In the post-lab, sketch the inside of the steel box showing how the wand and plates appear when fully inserted, and label all parts.
3. With nothing between the wand and the source, take 10 readings and record them in the post-lab.
4. Insert the 0.12 mm aluminum plate between the wand and the source, take 10 readings, and record them in the post-lab.
5. Remove the thin plate, insert the 0.60 mm aluminum plate, take 10 readings, and record them in the post-lab.
6. Return to your sketch and mark where you think the radioactive source is located inside the box. Label the source.
7. Work on the post-lab questions.

Name: ______________________ **Date:** ______________________

Collaborators: ______________________ **Lab Section:** ______________________

Postlab Exercises (24 points)

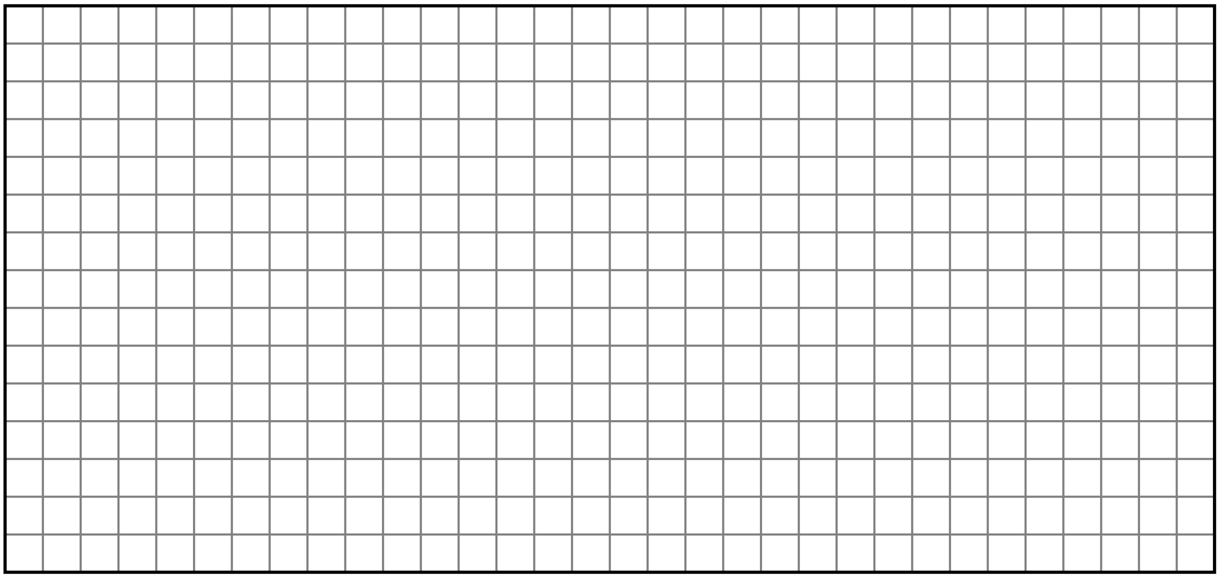
Submit the postlab to the TA at the end of the lab.

Experiment 1 (12 points)

Question 4 **3 points**

For each h_i recorded during the experiment, calculate the corresponding t_i in milliseconds and enter it in the table below. Pay attention to units.

Trial	Your Data		Partner's Data		Trial	Your Data		Partner's Data	
	h [cm]	t [ms]	h	t		h [cm]	t [ms]	h	t
1					14				
2					15				
3					16				
4					17				
5					18				
6					19				
7					20				
8					21				
9					22				
10					23				
11					24				
12					25				
13									



Question 5 **3 points**

Use the graph above to draw cumulative distribution functions for both your and your partner's reaction times. Choose an appropriate bin size based on your data. Include axes, labels, units, and a title. In future plots you will not be reminded to do this.

Question 6 **3 points**

Calculate the mean and standard deviation for each of the two data sets you collected. Do not forget units. Write your results in the form $\hat{\mu} \pm \hat{\sigma}$:

	Your Values	Partner's Values
Height	$\pm$	$\pm$
Time	$\pm$	$\pm$

Question 7 **3 points**

Use error propagation to compute the variance in your t measurements. How does it compare to the standard deviation you computed in the previous question?

Experiment 2 (12 points)

Question 8 **2 points**

Record your Geiger counter readings in the table below and sketch the inside of the steel box.

Reading	No Plate	0.12 mm Plate	0.60 mm Plate
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
Total			

Sketch of the inside of the steel box:

Question 9 **2 points**

Consider only your first reading from each of the three setups. What is the statistical error in each case? Does this single reading distinguish between the three setups? Explain.

Question 10 **2 points**

Sum all ten readings for each setup and enter the totals in the last row of your table. What is the statistical error in each total? Does the combined data distinguish between the three setups? Explain.

Question 11 **2 points**

For each setup, count how many of your ten readings fall within one error bar of the mean, and how many fall within two error bars. Organize your results in a table.

Question 12 **3 points**

Compare the error from a single reading to the error from the total of ten readings. Is their ratio what you expect? Explain. What errors would you expect for 100 readings? For 1000?

Question 13 **1 point**

How many readings would you need to unambiguously distinguish between the no-plate and 0.12 mm-plate setups? Explain your reasoning.