

Course and Section _____

Names _____

Date _____

ERROR ANALYSIS SIMULATION

Introduction

In this online experiment we are interested in learning how to treat data. The main point here is that any time a scientist measures a number; he/she must be able to tell his colleagues the accuracy of that number. Otherwise its not possible to tell whether the number agrees with the predictions of a theory; or for another scientist to check the experimental results.

Submit your answers using Blackboard.

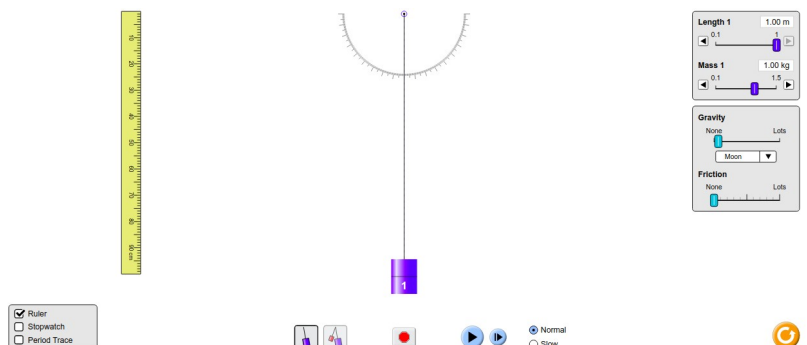
1 – Preliminary Questions

1. Suppose you measure the circumference of a circle and its diameter, then divide the circumference by the diameter. The result ought to be π . If your result is 3.05 did you just prove that π is 3.05?
2. Suppose you want to find the length of a rod. Does increasing the number of times you take a measurement of its length would improve the precision of your result?
3. You take thirty measurements in order to find the length of a rod, you calculate the average and obtain the value L_0 . Supposed that now a different person takes thirty measurements as well of the same rod, then calculates the average and obtains the same L_0 as you. Does it mean that both of you obtain a result with the same precision and accuracy?

2 – Pendulum

Open the simulation below and select Intro

(https://phet.colorado.edu/sims/html/pendulum-lab/latest/pendulum-lab_en.html)



On the top right corner select set $Length = 0.8$ m and $Gravity = Moon$. Use the cursor to drag the pendulum away about 20° from its equilibrium position (at the bottom). Then press *play*, the pendulum will start to oscillate. Your goal is to find its period of oscillation T (how long it takes to get back to the initial position). To do so select the *Stopwatch*. You are not expected to measure the perfect 'true' value, instead the best estimate of the 'true' value is given by the average over N measurements

$$\bar{T} = \frac{1}{N} \cdot \sum_{i=1}^N T_i$$

Measure T ten times and keep a record of your data (on a piece of paper for example).

4. What is the average value \bar{T} that you have found? Use three digits

Next, we need to find a way to quantify how inaccurate were your measurements. First let's introduce d , the deviation from the average value

$$d_i = T_i - \bar{T}$$

Calculate d_i for each of your measurements T_i and keep a record of your data. Now calculate the average of \bar{d}

$$\bar{d} = \frac{1}{N} \cdot \sum_{i=1}^N d_i$$

5. What is the average value \bar{d} that you have found? Use three digits

Clearly, the average of the deviations d_i cannot be used as the error estimate, since it gives us zero (or very close to zero, depending on how you have rounded your digits). This is because the deviations with positive sign are always canceled by the deviations with negative sign.

In order to evaluate the inaccuracy of T , indicated as σ_{average} , first we need the standard deviation σ_T which is an estimate of the statistical dispersion of your measurements T_i about the average \bar{T} . The standard deviation of the sample is

$$\sigma_T = \sqrt{\frac{1}{(N-1)} \cdot \sum_{i=1}^N d_i^2}$$

6. What is the standard deviation σ_T of your T_i ? (Use three digits)

σ_{average} which is the error of the measurement of \bar{T} is related to the standard deviation σ_T as

$$\sigma_{\text{average}} = \frac{\sigma_T}{\sqrt{N}}$$

7. What is σ_{average} of your measurement? Use three digits

In this way the measurement of the period is expressed as $T = \bar{T} \pm \sigma_{\text{average}}$

8. If you were to increase the number of measurements what happens to the standard deviation?

9. If you were to increase the number of measurements what happens to the error of the average?

10. Are your answers to question 9 and question 2 consistent?

3 – The Gaussian distribution

You can use online tools to calculate several statistical quantities, as for example here (<https://www.mathsisfun.com/data/standard-deviation-calculator.html>)

A physicist uses a particular device to take measurements of the length L of a quantity. The following table shows the $N = 22$ different values obtained for L (in centimeters)

5.96	7.36	8.77	8.14	9.88	8.63	9.70	7.97	8.30	5.13	8.08
7.36	5.57	8.53	7.38	9.19	8.36	7.77	5.52	6.51	8.49	6.79

11. Calculate the average value of the length.
12. Calculate the standard deviation of the length. (Use the *sample* standard deviation not the *population* standard deviation).

The distribution of the possible measurements of L is supposed to follow a Gaussian distribution, that is a plot where the x -axis contains the values of L and the y -axis indicates how frequent that value was obtained during the measurement. Use this simulation to visualize the distribution of the values of L (http://digitalfirst.bfwpub.com/stats_applet/stats_applet_7_norm.html)

Input your average value of L and relative standard deviation.

Two numbers are displayed at the top of the green dots, report those two green numbers

13. Lower green number is
14. Bigger green number is

The numbers $0.1587 \approx 0.16$ at the bottom indicates the number of measurements with a value not contained in the range given by the two green numbers. Since the plot is symmetric that means $0.16 \times 2 = 0.32$ or 32% of the measurements have a value outside the range defined by the two green numbers. By using the language of statistics those values “are outside one standard deviation”

15. How many of the given 22 values are expected to be outside one standard deviation? (calculate the 32% of 22).
16. How many measurements have a value below one standard deviation?
17. How many measurements have a value above one standard deviation?
18. Are your answers to previous two questions in agreement with the distribution to be a Gaussian?