

✓ Tutorial 3: Distinguishability

Developed by Rebeckah Fussell for Cornell Physics Labs. Adapted by Cristina Schlesier for Spring 2025.

In this tutorial we will continue to apply the concepts of mean, standard deviation, and standard uncertainty of the mean. We will also learn about how to compare the means of different data sets and test for distinguishability.

Before we get started, let's load some packages and define the functions for standard deviation and standard uncertainty of the mean. These functions are identical to the ones in the previous tutorial.

Make sure to run the cell below before proceeding (SHIFT + ENTER). As always, remember code cells build on each other, so run each in order.

```
1 import numpy as np
2
3 %matplotlib notebook
4 %matplotlib inline
5 import matplotlib
6 #matplotlib.use('module://ipynb.backend_nbagg')
7
8 import matplotlib.pyplot as plt
9
10 #%matplotlib widget
11 from ipywidgets import *
12
13 def standard_deviation(data):
14     N = len(data)
15     return np.sqrt(np.sum((data - np.mean(data))**2)/(N-1))
16
17 def standard_unc_of_mean(data):
18     N = len(data)
19     return standard_deviation(data) / np.sqrt(N)
```

In the uncertainty tutorial we talked about taking repeated measurements of a phenomenon, and that we expect to see a distribution of values around the mean.

Often we need to determine if the means of two data sets of measurements are coming from the same underlying phenomenon. It is possible that we are measuring two different phenomena. It is also possible we are measuring the same phenomena. However, even if we are measuring the same phenomena when we take two different data sets, we would still expect to see some difference in the means of two different data sets.

For example, if we take several measurements of the period of a pendulum when released from 10 degrees, and several more measurements of the period at 20 degrees, we need a method to know if the average period is the same (and therefore we are measuring the same physical phenomenon) or different when released from the two angles. We are going to call this **distinguishability**.

This method should:

1. Take into account the difference between the means of the two data sets, such that a larger difference indicates the values are more distinguishable.
2. Take into account the uncertainty of the means of the two data sets, such that the difference in the means is balanced by how well we know the measurements.
3. Not depend on units so we can interpret the value without context.

To make comparisons between two measurements with uncertainties $A \pm \delta A$ and $B \pm \delta B$, we use a quantity known as t' . This is defined as:

$$t' = \frac{|A - B|}{\sqrt{\delta A^2 + \delta B^2}}$$

This ratio tells you how different two measurements are as a fraction of the uncertainty in the measurements. You may have heard particle physicists refer to 3-sigma or [5-sigma](#) effects: *sigma* here refers to the uncertainty, so a 3-sigma effect corresponds to a $t' = 3$.

For any experiment, A and B might be the means for two sets of repeated measurements (such as our means of repeated measurements of the period of pendula released from 10 and 20 degrees), or they might be individual measurements (such as the counts registered in a detector at two different elevations). δ should be read as "uncertainty in", so δA and δB are the uncertainties in the measurements A and B (*not* the uncertainty times the measurements).

Q. Explain how the formula for t' takes into account the three properties above.

[Your response here]

- 1.
- 2.
- 3.

Q. a) Say we measure $A = 2 \pm 1$ s and $B = 8 \pm 1$ s. Intuitively, would you say that A and B are distinguishable or indistinguishable? b) What about if the uncertainties were larger: are $A = 2 \pm 5$ s and $B = 8 \pm 5$ s distinguishable?

[Your response here]

- a)
- b)

Run the code below to calculate the t' value for these two sets of measurements.

```
1 def t_prime(A, dA, B, dB):
2     return abs(A - B) / np.sqrt(dA**2 + dB**2)
3
4
5 # an example of using t_prime
6 print('t-prime for a) measurements:', t_prime(2,1,8,1))
7
8 # second example
9 print('t-prime for b) measurements:', t_prime(2,5,8,5))
```

✓ Exploring t'

You can think of t' as telling us how far apart the means of the distributions are in "units" of uncertainty. The larger the t' value, the less likely it is that the two sets of measurements came from the same distribution. Below, we have two sets of measurements (one in blue and one in orange), with properties you can control in the code cell below. They have a relative mean (m), a number of measurements (N) and each a standard deviation.

Run the code below, and see what happens for different values of m , N , σ_1 , and σ_2 .

```
1 x=np.linspace(-5, 10, 1000)
2
3 # **** Adjust these parameters to explore how they impact t'. **** #
4 # **** The figure has a fixed x-range of -2,2 so keep the means **** #
5 # **** within those bounds. **** #
6 m=0.25      # Mean of the orange data relative to the mean of the blue data
7 N=20        # Number of measurements in each dataset
8 stdev1=.5   # Standard deviation of the blue dataset
9 stdev2=.5   # Standard deviation of the orange dataset
10 # **** #
11
12 fig, ax=plt.subplots(1,1, figsize=(5,5))
13 ax.cla()
14 x=np.linspace(-2, 2, 1000)
15 ax.fill_between(x,np.e**(-x**2/stdev1**2),alpha=.5)
16 ax.fill_between(x,np.e**(-(x-m)**2/stdev2**2), alpha=.5)
17 ax.errorbar(0, 1, xerr=stdev1/np.sqrt(N), capsize=5)
18 ax.errorbar(m, 1, xerr=stdev2/np.sqrt(N), capsize=5)
19 ax.set_title("t'={} ".format(round(np.sqrt(N)*m/np.sqrt(stdev1**2+stdev2**2), 3)))
20 fig.canvas.draw_idle()
21 fig.canvas.draw_idle()
22 plt.tight_layout()
23
```

Q. a) What combination of factors causes t' to increase? b) What combination of factors causes it to decrease? c) Does the difference in the mean have to change for t' to change?

[Your response here]

- a)
 - b)
 - c)
-

✓ Interpreting t' values

If we have two sets of measurements of the same physical phenomenon, we would expect to see a t' value of approximately 1 on average. This is because, on average, most (68%) of our measurements should be within one standard deviation of the mean¹ and so any two measurements should almost always be approximately one standard deviation away from each other.

After calculating t' for two measurements, you can evaluate their dissimilarity (or distinguishability) through the following interpretation:

1. $t' \approx 1$ If we have two sets of measurements and a t' value of approximately 1, then the sets are indistinguishable and they **may** represent the same physical phenomenon.
2. $t' << 1$ If we have a t' value of much less than 1, then it is possible that either we overestimated our uncertainties, or that our current level of precision is not good enough for the phenomenon that we are trying to measure.
3. $1 \lesssim t' < 3$ This is a grey area. It is still possible that our two sets of measurements are coming from the same phenomenon, but it is less likely than if our t' is somewhere close to 1.
4. $t' > 3$ If our t' is greater than 3, then it is unlikely that our two sets of measurements were measuring the same phenomena. This means that we have distinguished between two sets of physical phenomena.

NOTE: $|t'| \leq 1$ **does not** mean that A and B are the same. It only tells us that the given data cannot allow us to distinguish between the two sets. For example, if you do a better measurement and decrease the uncertainties, you might later uncover a difference between A and B. That is, poor precision may be hiding a subtle difference!

In all cases, your t' value will give you **evidence** of one interpretation over another, but it **does not "prove"** an interpretation². Further testing is almost always needed to continue building evidence in favor of whatever conclusions you may be lead to draw.

Q. Based on these interpretations and your exploration with the blue and orange data above, what do you think we should do next in each of these 4 scenarios?

1. $t' \approx 1$
2. $t' << 1$
3. $1 \lesssim t' < 3$
4. $t' > 3$

[Your response here]

- 1.
 - 2.
 - 3.
 - 4.
-

✓ Pendulum Example

Let's try an example problem. An experimenter measured the period of a pendulum at 10 degrees and 20 degrees using a simple stopwatch. They measured the time for the pendulum to swing for one single period and conducted 14 trials for each angle. The data they collected is given in the first lines of code below. The code then generates a histogram so we can visualize the distribution of our data.

Run the code cell below.

```
1 #enter pendulum period data
2 plt.figure()
3 ten_degs = np.array([1.23, 1.36, 1.35, 1.36, 1.30, 1.27, 1.30, 1.32, 1.26, 1.26, 1.38, 1.29, 1.29, 1.32])
4 twenty_degs = np.array([1.37, 1.35, 1.38, 1.27, 1.33, 1.33, 1.26, 1.36, 1.36, 1.27, 1.29, 1.44, 1.29, 1.32])
5
```

```

6 #plot distributions
7 plt.hist(ten_degs,bins = 7)
8 plt.title("Histogram of pendulum period at 10 degrees")
9 plt.ylabel('Counts')
10 plt.xlabel('Period s')
11
12 plt.figure();
13 plt.hist(twenty_degs,bins = 7)
14 plt.title("Histogram of pendulum period at 20 degrees")
15 plt.ylabel('Counts')
16 plt.xlabel('Period s')
17 plt.show()

```

Q. Calculate standard deviation of the 10 degree data set. *Hint: replace the dots with the name of the ten degree data array as it was defined in the previous cell.*

```
1 standard_deviation(...)
```

Q. Calculate standard deviation of the 20 degree data set.

```
1 #write your function call here:
```

Now we will shift to asking some questions about the experimenter's data set based on what you learned in class.

Q. Based on the data, what is the uncertainty in the measurements of the period from **instrumental precision**? *Hint: scroll back up to where the ten_degs and twenty_degs arrays are defined above. Notice that all the data reports the same number of digits after the decimal point. What does that say about the precision of the timer?*

[Your response here]

Q. Calculate the uncertainty in a single trial of each set of measurements for the period of this pendulum from: i) ten degrees and ii) twenty degrees. *Hint: you will need to call one of the functions defined above (standard_deviation(), standard_unc_of_mean(), or t_prime()). Fill the parentheses with the name of the data set you want to make a calculation on.*

```
1 #write function call here:
```

[Your response here]

Q. Determine the measured value of the period from each data set (ten degrees and twenty degrees). Express your answer as (best estimate) $\pm$ (uncertainty).

```

1 print("i) ten degrees:")
2
3 #edit two lines below
4 mean_10 = np.mean(...) # This function calculates the mean of a list of numbers
5 unc_10 = ...           # Assign this variable with the correct value for the uncertainty
6
7 print("mean is " + str(mean_10) + " +/- " + str(unc_10))

```

```

1 print("ii) twenty degrees:")
2
3 #edit two lines below
4 mean_20 = ...
5 unc_20 = ...
6
7 print("mean is " + str(mean_20) + " +/- " + str(unc_20))

```

Q. Quantitatively compare the measurements of the period of this pendulum when released from 10 degrees and 20 degrees. Use the t' function below and replace the ... with the relevant uncertainties.

```
1 print(t_prime(mean_10, ..., mean_20, ...))
```

Q. Interpret the results of your quantitative comparison in the previous question. Propose at least two next steps the experimenter could reasonably take based on these results.

[Your response here]

Interpretation:

Next step 1:

Next step 2:

Save your notebook with all your answers to the questions, modified code cells, and output from each code cell. Submit your notebook by saving it as a PDF and uploading the PDF to the Gradescope assignment titled **Tutorial 2**.

Footnotes

1 The 68% refers to the area under a Gaussian (or Normal) distribution. For counting events, such as in particle physics, measurements are better characterized by a Poisson distribution. Fortunately, the [central limit theorem](#) any distribution converges towards a Gaussian distribution with a large enough number of measurements. [back up](#)

2 We would advise you to avoid the words *proof* or *prove* in your discussions of your experiments. In science, nothing is ever proven definitely. Any implication otherwise by textbooks or popular writing on science is sloppy at best and misleading at worst. In science, we only have increased or decreased levels of confidence regarding a claim, by way of experimental evidence, but we should not expect to arrive at an absolute truth. [back up](#)