

Statistics in Science



By C. Kohn

Waterford Agricultural Sciences



How do we know we know?



- ❧ A major concern in science is proving that what we have observed would occur again if we repeated the experiment.
 - ❧ Random things can affect our experiments.
 - ❧ Your samples might be affected by little things that change or skew your results.
 - ❧ The trends you find in your experiment may not occur in a different experiment done in the same way.

- ❧ We must always be prepared to answer the Scientist's Questions:
 - How do I know I am not wrong?
 - How do I know that this will always occur every time I do this experiment?



Target Shooting & Statistics



- ❧ **Data in research is sort of like target practice.**
 - ❧ When you are shooting at a target, you want all of your shots to be close together.
 - ❧ The closer your shots are to each other, the better.
 - ❧ You also need to take a lot of shots in order to be accurate.
 - ❧ The more times you shoot at a target, the more accurate of a shooter you are.

- ❧ **Statistics are the same: the more data we have, and the more similar each number is to each other, the better.**



Source: topendsports.com

Science & Statistics



- ❧ **In science, we can use statistical equations to determine whether or not we can be confident in our results.**
 - ❧ In other words, the use of statistics can tell us whether our experimental results are reliable.
- ❧ **If we are likely to see similar results every single time, this means that our results are reliable.**
 - ❧ On the other hand, if we get very different results each time we do an experiment, our data is less reliable.

❧ The more variable our data, the less reliable it is.

Less reliable



❧ The less our data varies, the more reliable it is.

More reliable



The “Real” Average



- ❧ When we need to calculate the average of our data (or “mean”), we can encounter problems with reliability.
 - ❧ Mean: the numerical **average** of data (*mean = average*)
 - ❧ It is calculated by dividing the sum of the numbers by the sample size. Mean = (Sum of Data) / (Sample Size)
 - ❧ E.g. mean of 1,2,&3 would be $(1+2+3)/3 = 2$. Our mean is 2.

- ❧ When we take the average of something, we are using a number that can change as we gain or lose data.
 - ❧ For example, imagine if we wanted to know the mean height of this class.
 - ❧ To obtain this number, we would...
 - ❧ 1. Record each person’s height
 - ❧ 2. Add them all together, and
 - ❧ 3. Divide by the number of students we have to get the “mean height”.



Averages (cont.)



- ❧ **However, if we gained or lost a student, the mean (or *average*) height would change.**
 - ❧ The “average height” is not one number; it can change!

- ❧ **If our class did not have very many students, the addition of one more person’s height would have a big impact on the calculated average.**
 - ❧ On the other hand, if we had 1000 students in our class, the addition of one more person’s height would hardly change the calculated average.

- ❧ **If the new person’s height was very similar to the average, our calculated average would not change much.**
 - ❧ On the other hand, if they were 6’7”, our calculated average would change a lot more.

Factors that Affect Data Reliability



☞ Things that affect the *reliability* of our data include:

☞ How **similar** our data is:

☞ The more similar the data, the more reliable our average will be.

☞ E.g. if all of our students are between 5'10" and 6'1", we would have more reliable data than if the range of the data was greater (such as if the range was between 4'5" and 7'1")

☞ The **amount** of data we have:

☞ The more data we have, the more reliable our average will be.

☞ E.g. if you flip a coin 3 times, you might get 2 heads, 1 tail.

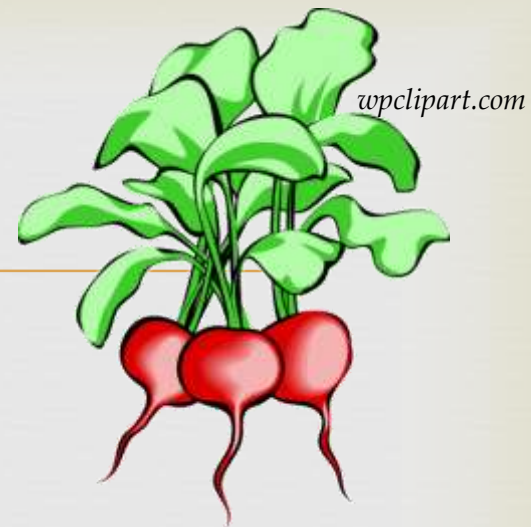
☞ If you flip a coin 10 times, you might get 6 heads and 4 tails.

☞ If you flip a coin 100 times, you might get 49 heads, 51 tails

☞ Each time we get closer to the "real" average of 50/50



Examples



- ❧ For example, let's imagine you want to know how UV light affects radish growth.
 - ❧ If you have only six plants, your data will not be very reliable.
 - ❧ If you have thousands of plants, your data will be much more reliable.
 - ❧ If the height of your plants varies a lot (e.g. some are 2 inches, some are 20 inches), then your data will not be very reliable.
 - ❧ If all your plants are almost the same size, your data will be very reliable.
- ❧ So *how* do we know for sure if our data is reliable or not?

Standard Deviation




☞ Standard Deviation is a measurement of how much our data varies.

☞ Low variance means your data is all very similar.

*These corn plants would have **low** SD* → 

☞ High variance means your data is very dissimilar.

*These corn plants would have **high** SD* → 

☞ Standard deviation is calculated by the following formula:

$$\text{Standard Deviation} = \sqrt{\frac{[(\text{data}_a - \text{avg})^2 + (\text{data}_b - \text{avg})^2 + \dots + (\text{data}_x - \text{avg})^2]}{(n-1)}}$$

*SD = stand. dev
n = sample size*

Standard Deviation Example



- ❧ For example, let's pretend that our radish heights were:
6.1 ; 5.8 ; 7.2 ; 4.3 ; 5.5 ; 5.8 cm
- ❧ The average (or mean) height would be:
 $(6.1 + 5.8 + 7.2 + 4.3 + 5.5 + 5.8)/6 = 34.7/6 = 5.8 \text{ cm}$
- ❧ To calculate standard deviation (s) we would subtract the mean value from each individual value, square it, divide by n-1, and take the square root:
- ❧ $\sqrt{[(6.1-5.8)^2 + (5.8-5.8)^2 + (7.2-5.8)^2 + (4.3-5.8)^2 + (5.5-5.8)^2 + (5.8-5.8)^2]/(6-1)} =$
- ❧ $\sqrt{[(0.3^2) + (0^2) + (1.4^2) + (-1.5^2) + (0.3^2) + (0^2)](5)} =$
- ❧ $\sqrt{[(0.09) + (0) + (1.96) + (2.25) + (0.09) + (0)] / 5} =$
- ❧ $\sqrt{4.39 / 5} = \sqrt{0.878} = 0.94 \text{ cm}$
- ❧ Our Standard Deviation score is 0.94 cm
 - ❧ (note: SD is measured in the same units as our data)

Standard Deviation



Source: istockphoto.com



- ❧ **Standard Deviation is a measure of variability.**
 - ❧ We want our data to be very similar; we don't want it to be spread out.
 - ❧ Data is like butter – we want it in a tight form, like a stick. We don't want it to melt and spread everywhere.

- ❧ **Standard Deviation is a measure of how varied your data is.**
 - ❧ However, as we said before, both variance and the size of your sample affect the reliability of your data.
 - ❧ Standard Deviation is only a measure of *variance*.

Standard Error



❧ Standard Error is a measurement of reliability of a data sample; it involves both the size of your data sample and the amount of variance of your data.

❧ Standard Error is calculated by dividing your Standard Deviation by the square root of your sample size.

❧

| | | |
|-----------------|--|-------------------------------|
| Standard Error= | $\frac{\text{Standard Deviation}}{\sqrt{n}}$ | $n = \text{your sample size}$ |
|-----------------|--|-------------------------------|

❧ Standard Error is a measure of the reliability of your data.

❧ It uses both the *size* of the data sample and the *variability* of the data.

Radish Standard Error Example



For example, for our hypothetical radishes:

Our 6 radish heights were:

6.1 ; 5.8 ; 7.2 ; 4.3 ; 5.5 ; 5.8 cm

Our mean was 5.8 cm.

Our Standard Deviation was 0.94 cm.

Our Standard Error is $0.94 / \sqrt{6} = 0.38$ cm

Key Benefits of Standard Error

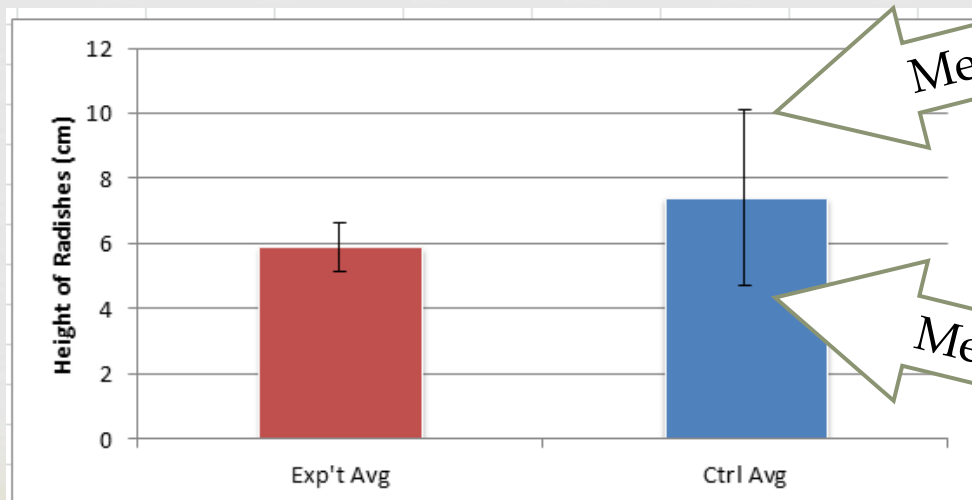


- ❧ **Standard Error tells us the likelihood of getting the same result if you repeated the experiment again.**
 - ❧ For example, if you are very likely to get the same average if you did the experiment again, you would have a small Standard Error.
 - ❧ If your results were more likely to be different, you would have a large Standard Error.
 - ❧ We always want to have as small of a Standard Error value as possible.

Standard Error and Confidence



- Standard Error is used to calculate your Margin of Error.
 - The margin of error shows us all of the possible outcomes of our experiment if we repeated it with a 95% accuracy.
 - Margin of Error is usually equal to 2x the Standard Error on either side of the mean (or average).
 - In other words, multiply your standard error by 2 and then add and subtract it from your mean.



Standard Error and Research



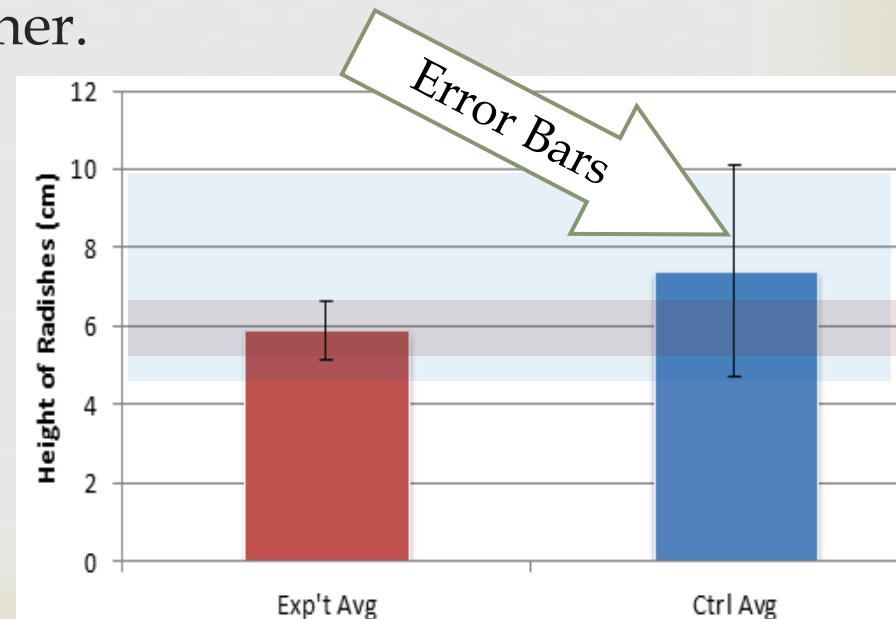
❧ **Standard Error is used to give us Error Bars.**

❧ Error bars are a visual depiction of your Margin of Error.

❧ If the error bars overlap, there is no statistical difference between the two groups (we have to treat them as if they are the same).

❧ E.g. these two groups are statistically the same because the error bars overlap with each other.

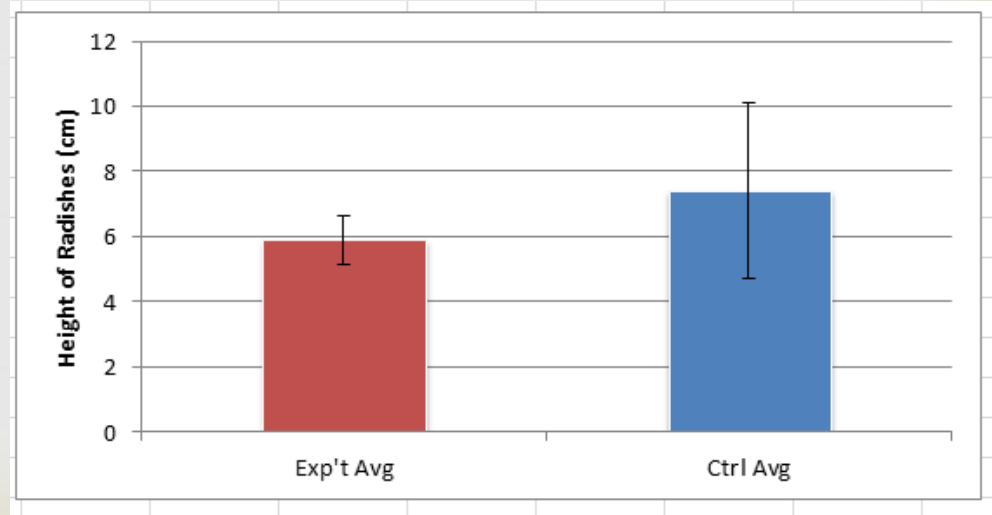
❧ Even though the blue bar is bigger, we have to treat them as if they are the same because statistically, they are.



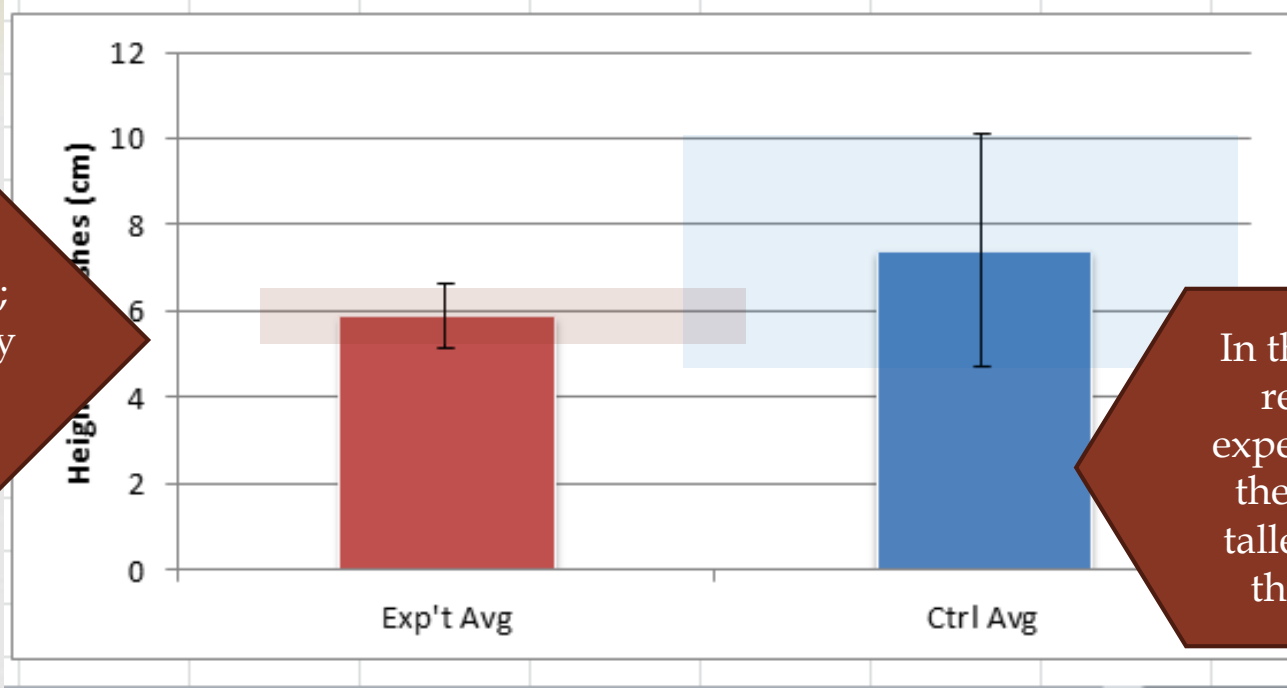
Error Bars



- ❧ In this example, the control has an average height (or *mean* height) that is over a full centimeter taller than the experimental average.
 - ❧ It looks as if the blue average is noticeably greater than the red average.
- ❧ However, the Error Bars (+/- 2 Standard Errors) overlap.
 - ❧ If your error bars overlap, this means that there is no statistically significant difference between the control and the experimental average.
 - ❧ You must treat them as if they are the same.

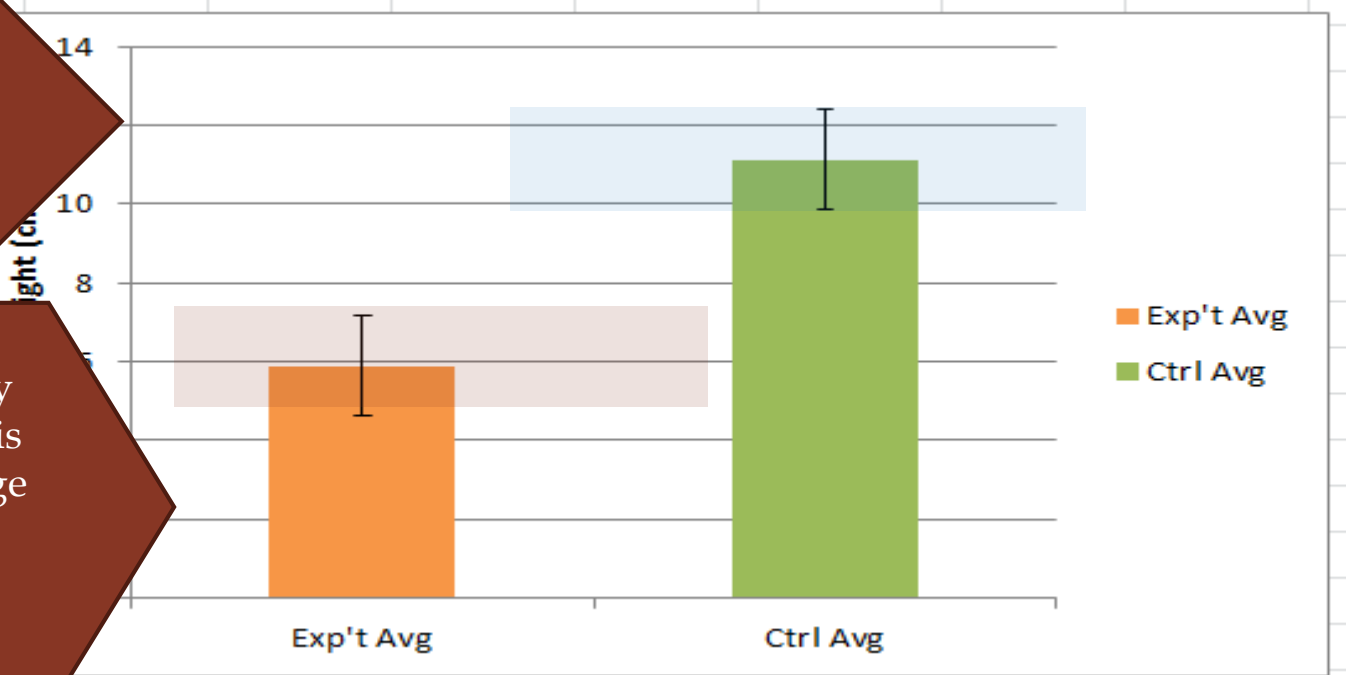


Error Bars overlap;
they are statistically
the same.



In this case, if we
repeated the
experiment again,
the red *could* be
taller on average
than the blue.

Error Bars do not
overlap; they are
statistically
different.



Exp't Avg
Ctrl Avg

No matter how many
times we repeated this
experiment, the orange
would always be
shorter on average
than the green.

Summary



- ❧ The more **consistent** the data, and the **larger** the sample size, the more reliable that data is.
 - ❧ Vice versa, small populations and highly variable data mean that it is less reliable.
- ❧ **Mean** is the *average* of the data (Sum of the Data / Sample Size)
- ❧ **Standard Deviation** is a measure of variability
- ❧ **Margin of Error** is the range in which we can be 95% sure of accuracy.
- ❧ **Standard Error** is a of measure the reliability of our data; it includes both variation *and* the sample size.
- ❧ **Error bars** can be made on graphs using $\pm 2x$ the Standard Error value.
 - ❧ Error bars indicate the range of accuracy of that data.
- ❧ If the **error bars** of two graphs overlap, those two graphs are considered statistically the same.
 - ❧ The error bars do not overlap, they are statistically different.

Calculating SD and SE



- ❧ Step 1: Calculate the average of each group
- ❧ Step 2: Subtract the average from each number
- ❧ Step 3: Square the result from Step 2 for each number
- ❧ Step 4: Add up each squared result
- ❧ Step 5: Divide the sum of each squared result by $(n-1)$
 - ❧ n = the number of numbers you have
- ❧ Step 6: Take the square root of your result from Step 5
 - ❧ This is your standard deviation
- ❧ Step 7: To find standard error, divide your calculated standard error by the square root of n

Review Concepts



- ❧ Definition of: 1) mean, 2) Standard Deviation, 3) Standard Error, 4) Error Bars
- ❧ Relationships between variability and accuracy
- ❧ 2 factors that increase the reliability of data
- ❧ How to tell if two graphs are statistically different
- ❧ How to calculate standard error and standard deviation.