Reporting Mean Values

Department of Agronomy, Iowa State University of Science and Technology

December 7, 2020

Variability

If you measure the height of five sorghum plants, they won't be exactly the same even if they have identical genotypes and are planted right next to each other. Subtle differences in solar radiation, soil water, soil nutrients, humidity, insect pressure (and other things) will result in small but measurable differences. Across a field there can be significant differences in soil properties, which will induce differences in yield. Across a state or region there will be significant differences in weather which will also effect yields. We call these differences *variability*. Agronomists need to understand what causes variability. In order to do that, we must be able to quantify it and determine how this variability affects data analysis.

The Standard Deviation

The standard deviation is the most common way to characterize variability. If a data set contains both large and small values, the standard deviation will be large. If all of the values are close together, the standard deviation will be small. Consider one characteristic of a sunflower phenotype, its stem diameter. (Perhaps breeding for larger stems would result in plants more resistant to lodging, which means falling over.) Say we had measurements of the width of several stems. If we were to add more measurements (from the same *population* or group of plants) to our set of data, the standard deviation would give us an idea of what values to expect. In this way the standard deviation tells us something about the uncertainty in the value of one measurement to the next.

The Normal Distribution

Variability in nature often can be described by what is called a *normal distribution*. A "distribution" is the different variations in a population that can be observed. For example, the distribution of sunflower stem diameter would be all the observed values of stem diameter for a particular sunflower genotype. An illustration of a normal distribution is shown in Figure 1. What makes a normal distribution special is its symmetric shape and useful mathematical qualities. In a normal distribution, more than two-thirds (68%) of values lie within plus–or–minus one standard deviation of the mean value, 95% lie within plus–or–minus two standard deviations, and essentially all values are within plus–or–minus three standard deviations.

Uncertainty in the Mean Value (The Standard Error)

The mean value of a data set is one way to represent all the measurements of the set. However, there will always be variability in the data (otherwise why calculate a mean value?). How does this

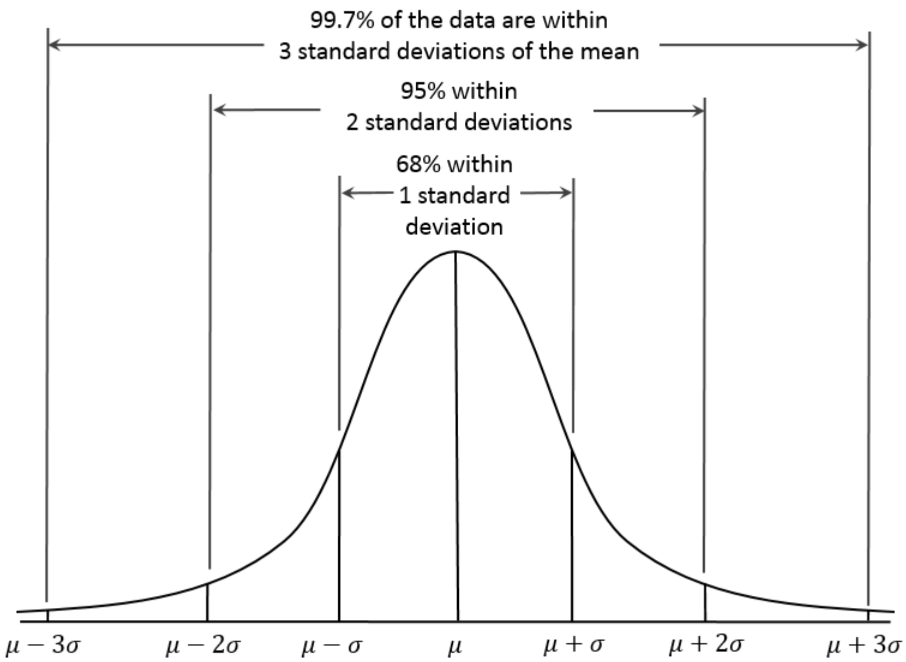

Figure 1: A normal distribution. The Greek letter μ represents the mean value of some characteristic (like stem diameter or plant height) of the population, while σ is the standard deviation, a measure of the variability in that characteristic. The height of the curve represents the number of individuals in the population that have each value of the characteristic. Credit: Dan Kernler, Wikipedia Commons.

variability affect the mean value? In other words, how well do we know the value of the mean? What is the *uncertainty* in the mean value? Think of it this way. If we measured every single stem in the population of sunflower planted in a field, we might be able to calculate the "true" mean value. But we seldom (if ever) have the time to do that! Consequently, the mean value of a data set is always just an estimate of the "true" mean value. We can, however, get an idea of how close the mean of our limited data set is to the "true" mean, and this information may be as important as the value of the mean itself. All data sets consists of a finite number of values that come from the distribution of all possible values that exist. We usually assume that the "real" population can be described by a normal distribution.

We can find the uncertainty in the mean value of a set of data by calculating what is called the *standard error*. The standard error is related to the standard deviation of the data set, which makes sense, but it also accounts for the size of the data set since the more data points we have, the better our estimate of the mean value should be.

> The uncertainty in the mean value of a data set can be characterized by the standard error of the data set, which is the standard deviation of the data set divided by the square root of the number of data points in the data set.

The smaller the standard error, the less uncertainty in the mean value.

Normally agronomists report a *95% confidence interval* by adding and subtracting twice the standard error to/from the mean value. A 95% confidence interval means that there is only a 5% chance that this interval around the mean value does not contain the "true" mean value.

Procedure for Reporting a Mean Value

When reporting a mean value, include its 95% confidence interval. Follow these directions.

1. Calculate the mean value of the data. Keep more decimal places than you will need in your final result.

2. Calculate the standard deviation of the data.

3. Calculate the standard error, which is the standard deviation of the data set, divided by the square root of the number of data points.

4. Double the standard error and then do both of these two things.

   (a) Round up the result to one significant digit.

   (b) Round up the result so that the place of the least significant digit matches the place of the least significant digit of the data.

   Use the smallest value, or the value that makes the most sense in this situation.

5. Round your mean value to the same place as the least significant digit in the value you ended up with in Step 4.

6. Report the mean value as your rounded mean plus or minus the value you found in Step 4. There is less than a 5% chance that the "true" mean is outside of this interval.

Is There a Significant Difference?

Say you have two sets of data and each set represents slightly different conditions. Perhaps you are evaluating yields in the same county but from different years. Or maybe one set is a collection of soil samples from one field and the other is a collection from another field. How can you tell if these sets of data are different or not? Well, you can compare their mean values, but a better way is to compare their mean values along with their respective uncertainties in their mean values. If we have 95% confidence that one mean value lies in one interval, and 95% confidence that the other mean value is within an interval that doesn't overlap with the interval around the mean value of the first set, then we will say that the two data sets are *significantly different*.

Here is an example. In 2016, students in AGRON 183 measured the stem diameter of two different sunflower genotypes, 31737 and 31859, which were grown in adjacent rows with the same management. For each genotype there were two replicate populations. The data are shown in Figure 2. For genotype 31737.1 (the first replicate of 31737) the mean value is 15.7355 mm and the standard deviation is 3.8031 mm. Since there are 29 data points, the standard error is $3.8031/\sqrt{29} =$ 0.7062 mm, and twice the standard error is 1.4124 mm. Rounding up 1.4124 mm to one significant digit gives 2 mm. Rounding up 1.4124 mm to the hundredths place (the place of the least significant digit of the data) gives 1.42 mm. 1.42 mm $<$ 2 mm so use 1.42 mm. Rounding the mean value to the least significant digit of this value results in 15.74 mm, and we report a mean value of the stem diameter of genotype 31737.1 to be 15.74 ± 1.42 mm. The $\pm$ sign indicates the true mean value is somewhere between $15.74 - 1.42 = 14.32$ mm and $15.74 + 1.42 = 17.16$ mm. (There is only a 5% chance that the mean value is *not* in this interval.) Table 1 displays the mean value for all four data sets in Figure 2.

The mean values of stem diameter and their respective uncertainties are graphically represented in Figure 3. Table 1 and Figure 3 show that *there is not a significant difference* in the mean values of stem diameter of the two replicates of 31737, because, in Figure 3, the mean value of 31737.1 is somewhere on that line, and the mean value of 31737.2 is somewhere on its respective line. Since these lines overlap, it is possible that the mean values of 31737.1 and 31737.2 could be the same. This makes sense: shouldn't replicates of the same genotype have the same mean value, if the environments (and management) of each replicate were essentially the same? (They were planted right next to each other.) There is also no significant difference in the mean values of stem diameter for the two replicates of 31859, nor is there between all four data sets since they all overlap (31737.1 and 31859.1 just barely overlap).

Summary

Using the standard error allows us to account for the effect of variability in the data as well as the finite number of our measurements on our estimate of the mean value. This allows us to best determine what the "true" mean value is. If we then want to compare to sets of data, we can determine whether different sets of data are significantly different or not depending on whether the 95% confidence interval for the mean value of one set overlaps with the 95% confidence interval of the other set.

There is one other thing to think about here... The standard deviation of a data set, and thus the standard error, represents both "natural" variability (caused by things like small differences in environmental conditions in our sunflower stem example) as well as the measurement uncertainty or precision of the measuring tool (the digital calipers). It is often possible to separate these two sources of variability if we make some assumptions. However, in many (but not all) cases relevant to agronomy the precision of the measuring tool is much smaller than the "natural" variability, such that the uncertainty introduced by the measurement instrument can be ignored. These types of assumptions may be important to carefully consider depending on the question you are trying to answer or the hypothesis you are testing.

Table 1: Mean values and 95% confidence intervals for the stem diameter of two sunflower genotypes (and two replicates of each genotype). The data is shown in Figure 2.

31737.1	31737.2	31859.1	31859.2
15.74 ± 1.42 mm	14.53 ± 2.15 mm	12.65 ± 1.83 mm	13.77 ± 1.52 mm

USDA North Central Regional Plant Introduction Center
09/01/16

stem diameter, mm

genotype	31737	31737	31859	31859
replicate	1	2	1	2
	14.66	17.75	20.11	14.35
	9.85	21.06	7.25	12.51
	16.61	13.56	10.65	12.07
	20.27	16.89	13.92	11.50
	17.66	12.94	15.72	13.79
	13.97	11.31	13.55	16.30
	17.26	9.13	10.00	8.34
	19.19	2.67	9.52	11.90
	13.87	8.33	14.46	15.75
	18.78	24.10	11.87	13.37
	9.16	15.90	5.74	16.34
	20.14	12.00	17.42	13.42
	16.05	12.34	9.22	26.71
	12.23	20.59	17.73	8.39
	14.41	11.67	7.19	16.86
	14.68	19.84	16.24	12.61
	19.46	18.23	15.72	16.54
	12.75	4.74	10.12	19.92
	15.94	12.80	12.60	11.35
	16.80	19.50	13.67	15.47
	16.75	11.66	10.96	10.01
	17.96	18.69	22.55	17.80
	15.19	15.12	4.20	3.19
	16.94	17.86	13.06	13.97
	8.22			14.88
	20.30			18.79
	11.79			16.63
	10.52			9.58
	24.92			8.69
				6.17
				16.79
				14.36
				10.83
				18.84

Figure 2: Stem diameter of two different genotypes, each with two replicates.

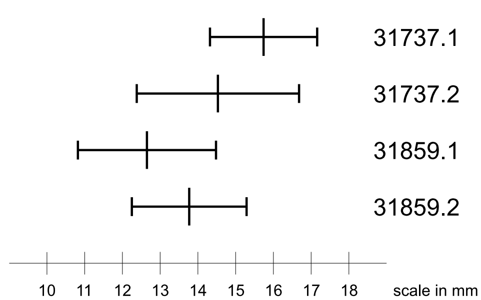

Figure 3: The mean values and 95% confidence intervals for the stem diameter of two sunflower genotypes (and two replicates of each genotype). This is a graphical representation of the information in Table 1.