How to Interpret Standard Deviation and Standard Error in Survey Research

Standard Deviation and Standard Error are perhaps the two least understood statistics commonly shown in data tables. The following article is intended to explain their meaning and provide additional insight on how they are used in data analysis. Both statistics are typically shown with the mean of a variable, and in a sense, they both speak about the mean. They are often referred to as the "standard deviation of the mean" and the "standard error of the mean." However, they are not interchangeable and represent very different concepts.

Standard Deviation

Standard Deviation (often abbreviated as "Std Dev" or "SD") provides an indication of how far the individual responses to a question vary or "deviate" from the mean. SD tells the researcher how spread out the responses are -- are they concentrated around the mean, or scattered far & wide? Did all of your respondents rate your product in the middle of your scale, or did some love it and some hate it?

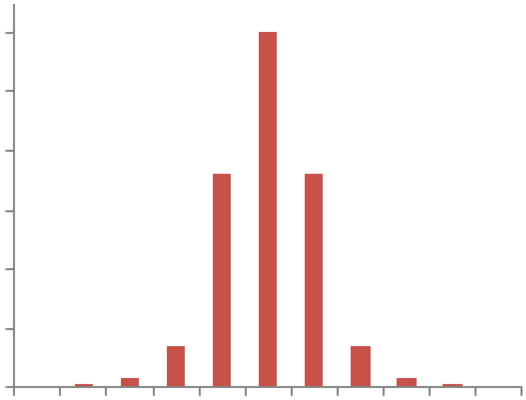
Let's say you've asked respondents to rate your product on a series of attributes on a 5-point scale. The mean for a group of ten respondents (labeled 'A' through 'J' below) for "good value for the money" was 3.2 with a SD of 0.4 and the mean for "product reliability" was 3.4 with a SD of 2.1. At first glance (looking at the means only) it would seem that reliability was rated higher than value. But the higher SD for reliability could indicate (as shown in the distribution below) that responses were very polarized, where most respondents had no reliability issues (rated the attribute a "5"), but a smaller, but important segment of respondents, had a reliability problem and rated the attribute "1". Looking at the mean alone tells only part of the story, yet all too often, this is what researchers focus on. The distribution of responses is important to consider and the SD provides a valuable descriptive measure of this.

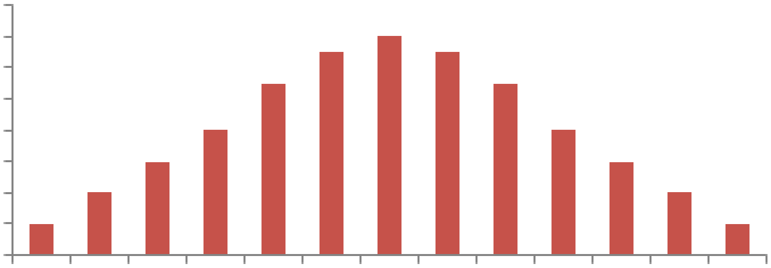
Respondent:	Good Value for the Money:	Product Reliability:
A	3	1
B	3	1
C	3	1
D	3	1
E	4	5
F	4	5
G	3	5
H	3	5
I	3	5
J	3	5
Mean	3.2	3.4
Std Dev	0.4	2.1

Two very different distributions of responses to a 5-point rating scale can yield the same mean. Consider the following example showing response values for two different ratings. In the first example (Rating "A") the Standard Deviation is zero because ALL responses were exactly the mean value. The individual responses did not deviate at all from the mean. In Rating "B", even though the group mean is the same (3.0) as the first distribution, the Standard Deviation is higher. The Standard Deviation of 1.15 shows that the individual responses, on average*, were a little over 1 point away from the mean.

Respondent:	Rating "A"	Rating "B"
A	3	1
B	3	2
C	3	2
D	3	3
E	3	3
F	3	3
G	3	3
H	3	4
I	3	4
J	3	5
Mean	3.0	3.0
Std Dev	0.00	1.15

Another way of looking at Standard Deviation is by plotting the distribution as a histogram of responses. A distribution with a low SD would display as a tall narrow shape, while a large SD would be indicated by a wider shape.


Histogram of distribution with small SD


Histogram of distribution with large SD

SD generally does not indicate "right or wrong" or "better or worse" -- a lower SD is not necessarily more desireable. It is used purely as a descriptive statistic. It describes the distribution in relation to the mean.

*Technical disclaimer: thinking of the Standard Deviation as an "average deviation" is an excellent way of conceptionally understanding its meaning. However, it is not actually calculated as an average (if it were, we would call it the "average deviation"). Instead, it is "standardized," a somewhat complex method of computing the value using the sum of the squares. For practical purposes, the computation is not important. Most tabulation programs, spreadsheets or other data management tools will calculate the SD for you. More important is to understand what the statistics convey.

Standard Error

The Standard Error ("Std Err" or "SE"), is an indication of the reliability of the mean. A small SE is an indication that the sample mean is a more accurate reflection of the actual population mean. A larger sample size will normally result in a smaller SE (while SD is not directly affected by sample size).

Most survey research involves drawing a sample from a population. We then make inferences about the population from the results obtained from that sample. If a second sample was drawn, the results probably won't exactly match the first sample. If the mean value for a rating attribute was 3.2 for one sample, it might be 3.4 for a second sample of the same size. If we were to draw an infinite number of samples (of equal size) from our population, we could display the observed means as a distribution. We could then calculate an average of all of our sample means. This mean would equal the true population mean. We can also calculate the Standard Deviation of the distribution of sample means. The Standard Deviation of this distribution of sample means is the Standard Error of each individual sample mean. Put another way, Standard Error is the Standard Deviation of the population mean.

Sample:	Mean
1st	3.2
2nd	3.4
3rd	3.3
4th	3.2
5th	3.1
.	.
.	.
.	.
Mean	3.3
Std Dev	0.13

Think about this. If the SD of this distribution helps us to understand how far a sample mean is from the true population mean, then we can use this to understand how accurate any individual sample mean is in relation to the true mean. That is the essence of the Standard Error. In actuality we have only drawn a single sample from our population, but we can use this result to provide an estimate of the reliability of our observed sample mean.

In fact, SE tells us that we can be 95% confident that our observed sample mean is plus or minus roughly 2 (actually 1.96) Standard Errors from the population mean.

The below table shows the distribution of responses from our first (and only) sample used for our research. The SE of 0.13, being relatively small, gives us an indication that our mean is relatively close to the true mean of our overall population. The margin of error (at 95% confidence) for our mean is (roughly) twice that value (+/- 0.26), telling us that the true mean is most likely between 2.94 and 3.46.

Respondent:	Rating:
A	3
B	3
C	3
D	3
E	4
F	4
G	3
H	3
I	3
J	3
Mean	3.2
Std Err	0.13

Summary

Many researchers fail to understand the distinction between Standard Deviation and Standard Error, even though they are commonly included in data analysis. While the actual calculations for Standard Deviation and Standard Error look very similar, they represent two very different, but complementary, measures. SD tells us about the shape of our distribution, how close the individual data values are from the mean value. SE tells us how close our sample mean is to the true mean of the overall population. Together, they help to provide a more complete picture than the mean alone can tell us.