Understanding the Difference Between Standard Deviation and Standard Error of the Mean, and Knowing When to Use Which

Chittaranjan Andrade[1]

Abstract

Many authors are unsure of whether to present the mean along with the standard deviation (SD) or along with the standard error of the mean (SEM). The SD is a descriptive statistic that estimates the scatter of values around the sample mean; hence, the SD describes the sample. In contrast, the SEM is an estimate of how close the sample mean is to the population mean; it is an intermediate term in the calculation of the 95% confidence interval around the mean, and (where applicable) statistical significance; the SEM does not describe the sample. Therefore, the mean should always be accompanied by the SD when describing the sample. There are many reasons why the SEM continues to be reported, and it is argued that none of these is justifiable. In fact, presentation of SEMs may mislead readers into believing that the sample data are more precise than they actually are. Given that the standard error is not presented for other parameters, such as difference between means or proportions, and difference between proportions, it is suggested that presentation of SEM values can be done away with, altogether.

Keywords: Standard deviation, standard error, standard error of the mean, confidence interval, statistical significance, graphs

Researchers who are knowledgeable about statistical tests are sometimes uncertain about the basics; few, for example, can correctly explain what the P value is.[1] In a similar vein, although most researchers know what the standard deviation (SD) and standard error of the mean (SEM) are, few can explain which should be used where and why. This article provides a simple clarification.

Standard Deviation

When we report our research, we need to describe our sample because the findings of our study can only be generalized to people who are similar to those whom we studied. We use descriptive statistics for this purpose. For quantitative variables, we report measures of central tendency and measures of dispersion. Measures of central tendency are the mean, median, and mode. Measures of dispersion are the range, SD, and interquartile range. It is as simple as that; we must report the SD as a measure of dispersion when we describe the sample, and the SEM does not come anywhere into the picture. This holds true whether we are describing the sample in numbers and words or in a figure.

What is the SD and why do we use it? If we regard distance from the mean as a positive number, the SD conceptually tells us how far from the mean the average person is. This indicates that if the SD is large, the values are widely scattered around the mean. In contrast, if the SD is small, the scatter is also small. Thus, the mean tells us what the average value is and the SD tells us what the average scatter of values is, around the mean. Taken together, especially along

ACCESS THIS ARTICLE ONLINE
Website: journals.sagepub.com/home/szj
DOI: 10.1177/0253717620933419

with the range, these statistics give us a good mental picture of the sample. Note, again, that the SEM does not come anywhere into this picture.

As an important aside, in a normal distribution there is a specific relationship between the mean and SD: mean ± 1 SD includes 68.3% of the population, mean ± 2 SD includes 95.5% of the population, and mean ± 3 SD includes 99.7% of the population. In this regard, published tables of area under the normal curve permit us to calculate the probability of finding a value at any distance from the mean when distance from the mean is expressed in terms of the SD. This is another use of the SD.

Standard Error of the Mean

The SEM is not a descriptive statistic. It tells us nothing about the sample. Therefore, it is illogical to state Mean (M) ± SEM when describing a sample; only M ± SD is correct. Then, when should the SEM be reported? A good answer could be "never" and the reason for this is that the SEM is best considered as an intermediate term in the calculation of 95% confidence interval (CI) and (where applicable) in the estimation of statistical significance.

What is the SEM? We know that the mean value that we obtain in our study is only an approximation of the mean value in the population. We also know that if we repeat our study a large number of times we will obtain different values for the mean, each time. The SEM is the SD of the means obtained in these different hypothetical studies. Thus, the SEM describes not our sample but the distribution of the means in these hypothetical studies. A large value for the SEM indicates that the means in the hypothetical studies are widely scattered. A small value for the SEM indicates that the means in the hypothetical studies are closely clustered. Therefore, the SEM is a measure of the precision of the study mean. If our study is an average study, the SEM is a measure of how far our study mean is from the population mean.

This implies that the SEM does convey some useful information to the reader. However, because the SEM is only an intermediate step in the calculation of the 95% CI, and because the 95% CI is, for various reasons, a preferred descriptor of the relationship between the study mean and the population mean,[2] it is better to report the 95% CI than the SEM.

As a point of interest, an increase in sample size makes it more likely that the sample is representative of the population, and hence that the sample mean is representative of the population mean. This is why, although an increase in sample size does not affect the value of the SD,[3,4] it does reduce the value of the SEM. The SD divided by the square root of the sample size gives us the value of the SEM.[5]

Concluding Notes

Whereas the SD describes the dispersion of data points in the sample, the SEM describes the precision of the study mean in the context of the population mean. The two concepts are so different that there is really no excuse for not knowing which value to report and where. So why do people continue to report the SEM along with the mean?

There are many "reasons." The commonest reason, especially in basic science reporting, is that others do it. This is inexcusable because it demonstrates a lack of application of mind. Another reason is that the SD is a simple concept, whereas the SEM, being more abstract, conveys an aura of high science, as is necessary in a scientific report. This is not acceptable, either, because statistics are furnished to explain, not impress.

The third reason is that the SEM is always smaller than the SD, and so when it is presented along with the mean, it makes the data appear more precise. This, as a reason, is deceitful. In fact, it is deceitful even when the reader fully understands what the SEM is. No reader will multiply the SEM with the square root of the sample size to get an idea of the SD of the sample. So, even the educated reader will read on, with an impression that the results in the sample are more precise than they actually are.

As a final reason why authors may use the SEM, because the SEM is smaller than the mean, when M ± SD data are presented in figures, SDs may take the error bars outside the box; in contrast, presenting M ± SEM data allows the figure to remain compact. This is unjustifiable because, as already explained, the reader can interpret M ± SD or M along with 95% CI; however, the reader has no theoretical framework to interpret M ± SEM as a sample descriptor because the pairing, as already explained, is illogical.

Some journals, now, explicitly require authors to present SDs, not SEMs.[3] Readers are referred to Streiner[5] and Altman and Bland[4] for a further discussion on the subject.

Finally, just as there is a standard error (SE) for the mean, there is an SE for the difference between means, an SE for a proportion, an SE for the difference between proportions, an SE for a correlation coefficient, and so on. Nobody reports the values for any SE other than the SEM; so why should the SEM ever be reported?

References

1. Andrade C. The P value and statistical significance: misunderstandings, explanations, challenges, and alternatives. Indian J Psychol Med 2019; 41(3): 210–215.
2. Andrade C. A primer on confidence intervals in psychopharmacology. J Clin Psychiatry 2015 Feb; 76(2): e228-31.
3. Bartko JJ. Rationale for reporting standard deviations rather than standard errors of the mean. Am J Psychiatry 1985; 142: 1060.
4. Altman DG and Bland JM. Standard deviations and standard errors. BMJ 2005; 331: 903.
5. Streiner DL. Maintaining standards: differences between the standard deviation and standard error, and when to use each. Can J Psychiatry 1996; 41: 498–502.