


On Using Standard Deviation or Standard Error of the Mean

Farrokh Habibzadeh^{1,2,3,4}, MD 

¹Independent Research Consultant, Shiraz, Iran;

²Past President, World Association of Medical Editors;

³Editorial Consultant, The Lancet;

⁴Associate Editor, Frontiers in Epidemiology

Correspondence:

Farrokh Habibzadeh, MD

Independent Research Consultant, Shiraz, Iran

Email: Farrokh.Habibzadeh@gmail.com

Received: 01 March 2025

Accepted: 02 April 2025

Abstract

The mean value is commonly used as a measure of central tendency. It is frequently reported along with either the standard deviation (SD) or the standard error of the mean (SEM). While the SD reflects the dispersion of the data in both the sample and population, SEM indicates the precision of the mean. SEM is not commonly used in reporting science; however, the 95% confidence interval, which is calculated based on SEM, is frequently reported in scientific literature.

Please cite this article as: Habibzadeh F. On Using Standard Deviation or Standard Error of the Mean. Iran J Med Sci. 2025;50(5):274-277. doi: 10.30476/ijms.2025.106283.4041.

Keywords • Medical writing • Statistics • Normal distribution • Confidence intervals

Introduction

Many scientists who have a good command of statistics know that to inform readers of the center of a normally distributed dataset, the mean of the data is commonly reported.^{1,2} A review of the biomedical literature reveals that the mean is frequently reported along with either the standard deviation (SD) or the standard error of the mean (SEM), as indices of data dispersion. SD is predominantly reported in clinical articles, and SEM in basic science papers. For years, there have been debates, even among editors,³ on which of these two indices, SD or SEM, is more appropriate to report.^{4,5} I would like to discuss this issue through a case study, and I hope this will ultimately help you identify which index to use.

Case Study

Suppose we wanted to determine fasting blood sugar (FBS) distribution in a population, such as those residing in Shiraz, southern Iran. To do so, assume that we measured FBS in 36 randomly selected individuals (Box 1) and found that the mean and SD of this dataset were 106 and 30 mg/dL, respectively. Visual examination of the distribution revealed that the data very likely followed a normal distribution (figure 1). Here, I do not elaborate on how to determine if the distribution of a dataset follows a normal distribution or not; for the time being, only accept that there are ways to test this hypothesis and assume that our data were normally distributed.^{6,7}

Box 1: Fasting blood sugar levels measured in 36 studied participants. The data are sorted.

66, 66, 67, 69, 70, 73, 79, 80, 80, 81, 84, 90, 92, 94, 96, 97, 104, 105, 106, 109, 112, 113, 114, 119, 120, 123, 127, 127, 129, 130, 136, 145, 146, 147, 168, 179

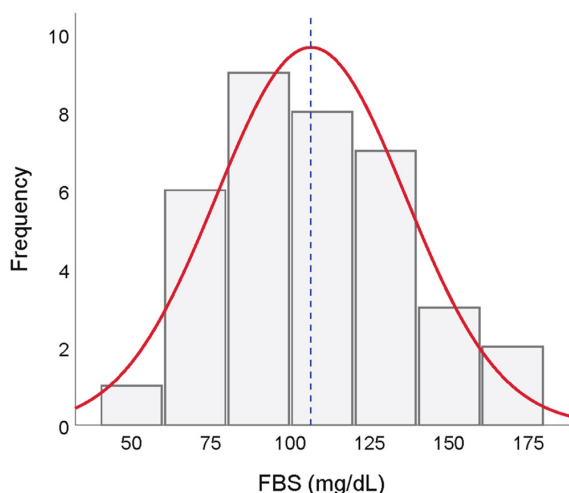


Figure 1: Histogram of the distribution of fasting blood sugar (FBS) values measured in 36 individuals. The dashed blue line represents the mean value of 106 mg/dL. The red curve represents the best normal curve fitted to the data. The graph was plotted using SPSS® for Windows® ver 26.

From basic statistics, we know that 95% of the data points that follow a normal distribution lie within two SDs around the mean.⁸ Therefore, we expect that 34 ($\approx 0.95 \times 36$ [sample size]) data points are within $106 \pm 2 \times 30$, from 46 to 166 mg/dL, which is consistent with our observations (Box 1, yellow figures). Therefore, the SD reflects the dispersion of data values around the mean.⁹

We conducted this study to estimate the mean FBS levels in the study population. We assumed that the calculated mean of 106 mg/dL was an acceptable estimate of the true population mean. This leads us to the important question of how accurate our measurement was. In other words, how far off is the calculated mean FBS from the true population mean?

Let us determine the answer to this question by conducting an imaginary experience.

What would be the mean FBS if we repeated the same study and measured the FBS in another group of 36 individuals randomly selected from the same population? With a high probability, the new value would be different from 106, the value we observed in our study (Box 1). If we draw a random sample of 36 people from this population and calculate the mean FBS for each sample, and then repeat the process 10 times, we will have 10 FBS means. The histogram representing the distribution of these 10 means obtained from our imaginary experience is not very informative (figure 2, leftmost panel). However, with an increase in the number of replicas (e.g., with 50, 200, and 10000 replicas [means]), the distribution of the FBS means approaches a normal distribution. The mean of this distribution (technically, the mean of the several mean values) is very close to the true population mean (figure 2, orange solid line).⁹

Similar to any other distribution, this distribution of FBS means also has an SD. In statistical parlance, this SD is termed the *standard error of the mean (SEM)*. Like all other normal distributions, we expect that 95% of the data points (in our example, mean values of study replicas) fall within the two SDs of this distribution (here, $2 \times \text{SEM}$) around the mean of the distribution (mean value of FBS means) (figure 2, interval bounded by the dashed red lines). For 10000 replicas (figure 2, rightmost panel), the mean value of the 10000 FBS means was 104.97 mg/dL; the SD of the 10000 FBS means (the SEM) was 5.01 mg/dL. Therefore, 9500 of 10000 replicas (95% of the FBS means)

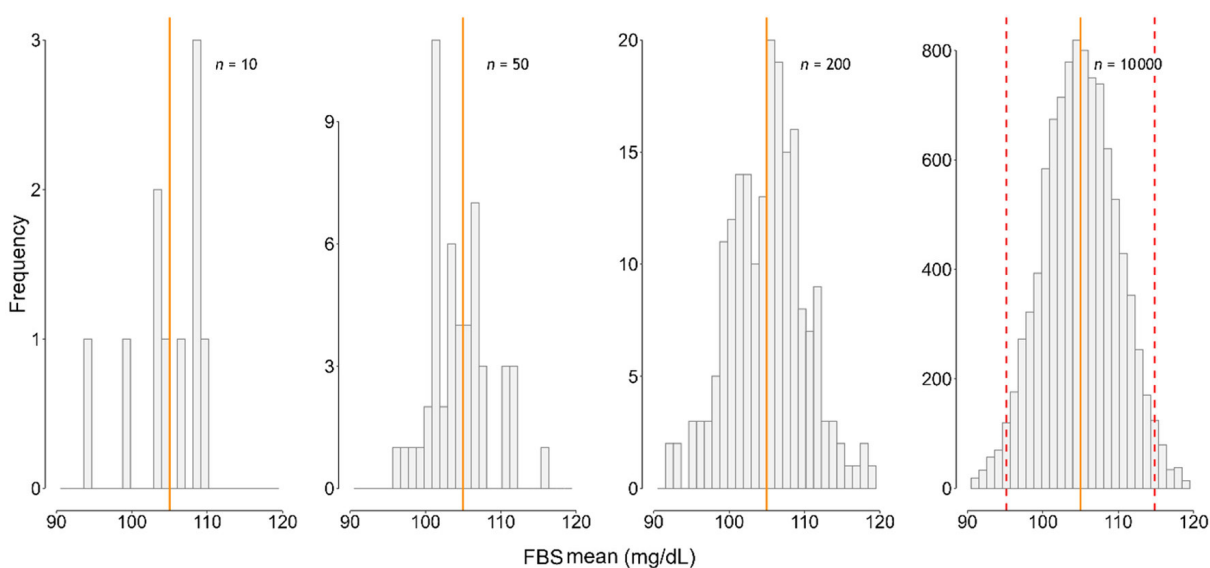


Figure 2: Histogram of the distribution of fasting blood sugar (FBS) means for four replica values (n of 10, 50, 200, and 10000 repeats). The orange lines represent the true population mean of 105 mg/dL. Dashed red lines indicate the 95% confidence interval for the mean ($\text{mean} \pm 2 \times \text{SEM}$).

should theoretically fall between 94.95 and 114.99 ($104.97 \pm 2 \times 5.01$) mg/dL. This interval is technically called the 95% confidence interval (CI) for the mean—the range of data values within which we can be 95% confident that the true population mean lies.

SD vs. SEM

Thus far, we have learned that while the SD describes the dispersion of data values around the mean, SEM reflects the dispersion of the distribution of *mean values* (if we conducted replicas of the study a large number of times) around the population mean, based on which we can calculate the 95% CI of the mean.^{1,7} While the SD reflects the dispersion of the data (in both the sample and population), SEM shows the precision of the mean. That is fine, nonetheless, should we necessarily conduct the same experience many times to calculate the SEM? The answer is [fortunately] *no*. It is possible to estimate SEM based on sample SD using the following equation:

$$SEM = \frac{SD}{\sqrt{n}} \quad (\text{Eq. 1})$$

where n represents the sample size (here, 36). Therefore, for our study, the estimated SEM is:

$$SEM = \frac{SD}{\sqrt{n}} = \frac{30}{\sqrt{36}} = 5 \text{ mg/dL} \quad (\text{Eq. 2})$$

which is very close to the value calculated from 10000 replicas, SEM of 5.01 mg/dL.

How to Report the Statistics

As examples, let us report the statistics necessary to answer the following three questions.

Q1. What are the mean and SD of the FBS distribution in the sample studied?

The mean (SD) FBS level in the study sample was 106 (30) mg/dL.

Q2. What are the mean and SD of the FBS distribution in the population?

We cannot be sure of this. However, based on the available evidence (our findings) the best estimate for the mean (SD) FBS of the study population is 106 (30) mg/dL.

Q3. How can we ensure that the mean of 106 mg/dL (what we observed) is close to the true population FBS mean?

Based on the calculated SEM (Eq. 2), we are 95% confident that the true population FBS mean lies between 96 and 116 ($\text{mean} \pm 2 \times \text{SEM}$, $106 \pm 2 \times 5$) mg/dL.

Therefore, we can state that “The mean (SD) FBS level was 106 (30) mg/dL.” or state that “The mean FBS level was 106 (95% CI 96 to 116) mg/dL.” depending on the message being conveyed. The former statement focuses on the distribution of data in the sample and population, whereas the latter stresses the precision of the measured mean value.

Examining Eq. 1 reveals that the SEM is always smaller than the SD.^{4,9} Many authors (and sometimes editors) chose to use SEM to pretend that the variability in their data was low. However, converting SEM into SD and *vice versa* is a piece of cake for you now that you are aware of their relationship (Eq. 1).

SEM is not commonly used in reporting science; the 95% CI, which is calculated based on SEM, is. In the charts and graphs (e.g., in bar charts), the error bars should preferably be the 95% CI rather than the SD or SEM. Finally, while you may now be familiar with the use of SD and SEM, it is important to remember that as an author, it is essential to adhere to the guidelines set by the journal to which you are submitting your work.

Conflict of Interest

Farrokh Habibzadeh, serving as an Editorial Board Member of the Journal, played no role in the handling of this manuscript at any stage. To ensure impartiality, the Editorial Board convened a team of independent experts to review the manuscript without his involvement or awareness.

References

- 1 Habibzadeh F. How to report the results of public health research. *Journal of Public Health and Emergency*. 2017;1. doi:10.21037/jphe.2017.12.02.
- 2 Lang TA, Altman DG. Basic Statistical Reporting for Articles Published in Biomedical Journals: The “Statistical Analyses and Methods in the Published Literature” or The SAMPL Guidelines. In: Smart P, Maison-neuve H, Polderman A, editors. *Science Editors’ Handbook*. European Association of Science Editors; 2013.
- 3 Habibzadeh F. Common Mistakes in Manuscripts Submitted to The IJOEM. *Int J Occup Environ Med*. 2018;9:61-2. PubMed PMID: 29667642.
- 4 Habibzadeh F. Common statistical mistakes in manuscripts submitted to biomedical journals. *European Science Editing*. 2013;39:92-4.

- 5 Habibzadeh F. Statistical Data Editing in Scientific Articles. *J Korean Med Sci.* 2017;32:1072-6. doi: 10.3346/jkms.2017.32.7.1072. PubMed PMID: 28581261; PubMed Central PMCID: PMC5461308.
- 6 Lilliefors HW. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association.* 1967;62:399-402. doi:10.1080/01621459.1967.10482916.
- 7 Habibzadeh F. Data Distribution: Normal or Abnormal? *J Korean Med Sci.* 2024;39:e35. doi: 10.3346/jkms.2024.39.e35. PubMed PMID: 38258367; PubMed Central PMCID: PMC10803211.
- 8 Altman DG, Bland JM. Statistics notes: the normal distribution. *BMJ.* 1995;310:298. doi: 10.1136/bmj.310.6975.298. PubMed PMID: 7866172; PubMed Central PMCID: PMC2548695.
- 9 Glantz SA. *Primer of Biostatistics.* 5th ed. New York: McGraw-Hill; 2002.