

Douglas Curran-Everett

Advan Physiol Educ 32:203-208, 2008. doi:10.1152/advan.90123.2008

You might find this additional information useful...

Supplemental material for this article can be found at:

<http://ajpadvan.physiology.org/cgi/content/full/32/3/203/DC1>

This article cites 31 articles, 29 of which you can access free at:

<http://ajpadvan.physiology.org/cgi/content/full/32/3/203#BIBL>

This article has been cited by 3 other HighWire hosted articles:

Explorations in statistics: hypothesis tests and P values

D. Curran-Everett

Advan Physiol Educ, June 1, 2009; 33 (2): 81-86.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Explorations in statistics: confidence intervals

D. Curran-Everett

Advan Physiol Educ, June 1, 2009; 33 (2): 87-90.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Reply to B. Kay

D. Curran-Everett and D. J. Benos

Advan Physiol Educ, December 1, 2008; 32 (4): 335-335.

[\[Full Text\]](#) [\[PDF\]](#)

Updated information and services including high-resolution figures, can be found at:

<http://ajpadvan.physiology.org/cgi/content/full/32/3/203>

Additional material and information about *Advances in Physiology Education* can be found at:

<http://www.the-aps.org/publications/advan>

This information is current as of August 27, 2009 .

Explorations in statistics: standard deviations and standard errors

Douglas Curran-Everett

Division of Biostatistics and Bioinformatics, National Jewish Health, and Department of Preventive Medicine and Biometrics and Department of Physiology and Biophysics, School of Medicine, University of Colorado Denver, Denver, Colorado

Submitted 21 March 2008; accepted in final form 6 May 2008

Curran-Everett D. Explorations in statistics: standard deviations and standard errors. *Adv Physiol Educ* 32: 203–208, 2008; doi: 10.1152/advan.90123.2008.—Learning about statistics is a lot like learning about science: the learning is more meaningful if you can actively explore. This series in *Advances in Physiology Education* provides an opportunity to do just that: we will investigate basic concepts in statistics using the free software package R. Because this series uses R solely as a vehicle with which to explore basic concepts in statistics, I provide the requisite R commands. In this inaugural paper we explore the essential distinction between standard deviation and standard error: a standard deviation estimates the variability among sample observations whereas a standard error of the mean estimates the variability among theoretical sample means. If we fail to report the standard deviation, then we fail to fully report our data. Because it incorporates information about sample size, the standard error of the mean is a misguided estimate of variability among observations. Instead, the standard error of the mean provides an estimate of the uncertainty of the true value of the population mean.

R; software; uncertainty; variability

IN 2004, Dale Benos and I proposed concise guidelines for reporting statistics (12).¹ We included a brief explanation or example to clarify each guideline, and we provided additional resources readers could use in concert with the guidelines. Nevertheless, we recognized that the guidelines—even with explanations and examples—were inadequate for readers who wanted to learn about fundamental concepts in statistics. Why are fundamental concepts in statistics important? They form the cornerstone of scientific inquiry. If we fail to understand these fundamental concepts, then the scientific conclusions we reach are more likely to be wrong. And wrong conclusions based on faulty reasoning is shoddy science (14).

Two series have been published in an effort to help readers learn about statistics: Altman's "Statistics and Ethics in Medical Research" (1–7) in the *British Medical Journal* and Healy's "Statistics from the Inside" (18–33) in the *Archives of Disease in Childhood*. These series are helpful, but, as you can imagine, they are entirely didactic. What do you expect, right? The problem is that most of us learn statistics like we learn science: by doing it.

When I teach and write about statistics, I want to engage my audience. To do this I use simulations as thought experiments my audience can see (11–14). From my perspective, the only thing better would be if my audience could run the simulations on their own. This series in *Advances in Physiology Education* provides an opportunity to do just that: we will investigate basic aspects of

statistics using a free software package. Among the concepts we will examine in future installments are the concepts behind *P* values and confidence intervals. My goals are to provide a theoretical framework for and a vehicle with which to illustrate each concept.

In this inaugural paper we explore the distinction between standard deviation and standard error, a distinction that has been discussed already (8, 10, 12–14, 16, 17). Before we begin that exploration, however, we need to learn a little about the software we will use to help us learn about concepts in statistics.

R: Software to Explore Concepts

In my statistics course, I use the freeware package R (34). R is a system—a language and an environment—for statistical analysis and data graphics. The R environment is a command-line environment in which `>` represents the command line. You can submit an R command in two ways: type the command in the interactive R Console, or submit the command from a script.² Because this series relies on R merely to explore fundamental concepts in statistics, I provide a script of the requisite R commands.

Regardless of whether you use a Mac or a PC, there are three preliminary steps to perform:

1. On your Desktop, create a folder called Advances.
2. Download the script `Advances_Statistics_Code.R` to your Advances folder.³
3. Download and install R.

Installation and basic operations. If you use a Mac, download R from

<http://cran.us.r-project.org/bin/macosx/> .

After you have installed R, double-click on `Advances_Statistics_Code.R` to open it. To submit particular commands in `Advances_Statistics_Code.R`, highlight the commands you want to submit and then press `⌘ ↵` (command key+enter). In the R Console, default colors are blue for command and black for result. Data graphics are shown in the Quartz device window.

If you use a PC, download R from

<http://cran.us.r-project.org/bin/windows/base/> .

After you have installed R, a shortcut for R will exist on your Desktop. To simplify the process of starting R from within your Advances folder, move this shortcut into your Advances

Address for reprint requests and other correspondence: D. Curran-Everett, Div. of Biostatistics and Bioinformatics, M222, National Jewish Health, 1400 Jackson St., Denver, CO 80206 (e-mail: EverettD@njc.org).

¹ These guidelines can be accessed through the American Physiological Society "Information for Authors" (9).

² A script is a file composed of R commands.

³ This file is available through the Supplemental Material link for this article on the *Advances in Physiology Education* website.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

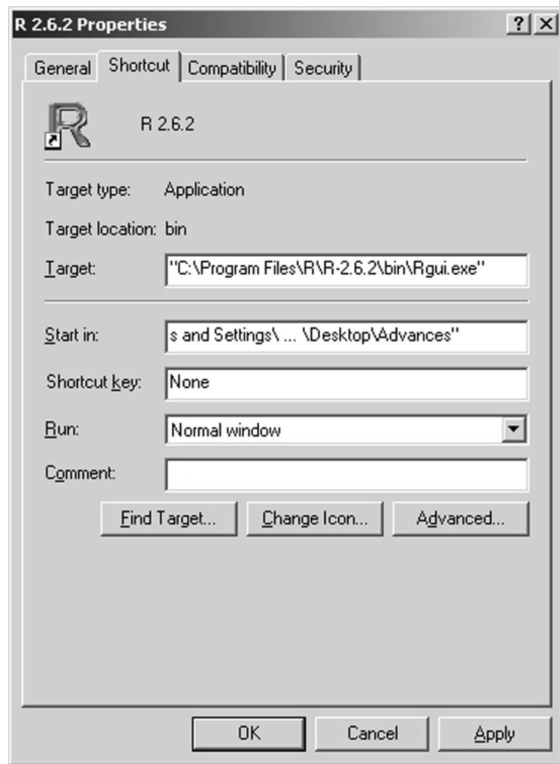



Fig. 1. Windows menu to change properties of the R shortcut so that R opens in C:\Documents and Settings\... \Desktop\Advances.

folder, right-click on the shortcut, and then click on Properties. Paste the full address (path) of your Advances folder

C:\Documents and Settings\... \Desktop\Advances

into the Start in: location (Fig. 1) and then click OK. Now double-click on the R shortcut to open R. To open Advances_Statistics_Code.R, click File|Open script...⁴ or click the Open script icon , select the script filename, and then click Open. Advances_Statistics_Code.R will open in the R Editor.

To submit particular commands in Advances_Statistics_Code.R, highlight the commands you want to submit, right-click, and then click Run line or selection. Or, after you highlight the commands you want to submit, you can simply type Ctrl+R. In the R Console, default colors are red for command and blue for result. Data graphics are shown in the R Graphics device window.

Basic syntax. The script Advances_Statistics_Code.R contains comments in addition to commands. Comments define sections of the script and explain many of the commands. The character # denotes a comment: all text after the first # on a line is a comment. For example, lines 5–8 of Advances_Statistics_Code.R are

```
# --Define population parameters and sample numbers-----
#
PopMean <- 0 # population mean
PopSD <- 1 # population standard deviation
```

The two commands in these lines of code, PopMean <- 0 and

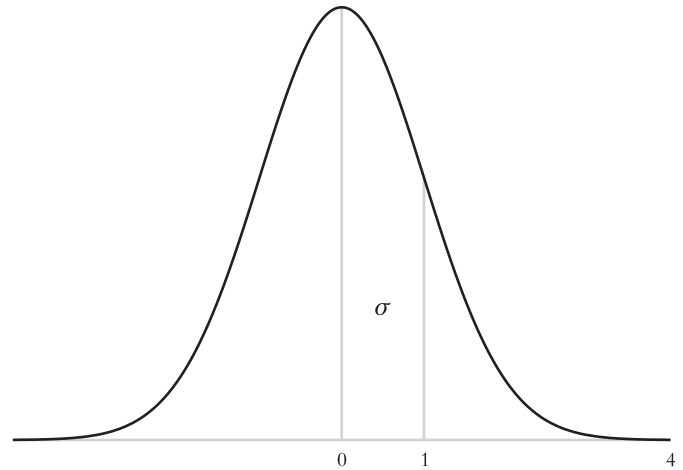


Fig. 2. The population. We can generate a standard normal distribution by transforming some random variable Y to z by the relationship $z = (Y - \mu)/\sigma$, where z represents the number of standard deviations Y is from the mean μ .

Table 1. Sample statistics calculated for each random sample

Column	Heading	Sample Statistic
1	Sample	Sample number
2	Mean	Mean \bar{y}
3	SD	Standard deviation s
4	SE	Standard error of the mean $SE\{\bar{y}\} = s/\sqrt{n}$
5	t	Observed value of $t = \bar{y}/SE\{\bar{y}\}$
6	LCI	Lower confidence interval bound
7	UCI	Upper confidence interval bound

PopSD <- 1, assign values to the variables PopMean and PopSD, the population mean and standard deviation.

If you highlight and then submit these lines of code, this is what you see in the R Console:

```
> # --Define population parameters and sample numbers-----
> #
> PopMean <- 0 # population mean
> PopSD <- 1 # population standard deviation
>
```

It is not obvious the commands have done a thing. If you type and then submit each variable name⁵ in the R Console, however, you see that the commands have assigned the values of 0 and 1 to PopMean and PopSD:

```
> PopMean
[1] 0
> PopSD
[1] 1
>
```

The Simulation: Observations and Sample Statistics

If we want to explore the distinction between standard deviation and standard error, we need some data. When I teach a class on regression, I introduce a data set in this way:

⁴ The notation *click A|B* means *click A*, then *click B*.

⁵ To do this, click within the R Console, type the variable name, and then press Enter.

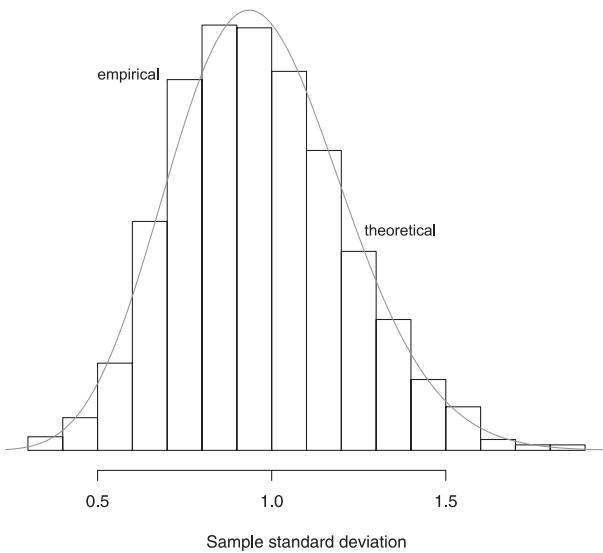


Fig. 3. Empirical (black) and theoretical (gray) distributions of the sample standard deviation for 9 observations. The empirical distribution is composed of 1000 sample standard deviations. The theoretical distribution of the sample standard deviation calculate using Eq. A1 in the APPENDIX. The commands in lines 76–87 of *Advances_Statistics_Code.R* calculate the theoretical distribution of the sample standard deviation, and the commands in lines 89–98 create this data graphic. To generate this data graphic, highlight and submit the lines of code from Figure 3: first line to Figure 3: last line.

It is difficult to choose an example that is relevant to everyone. So instead, I want to use an example that is relevant to no one: cement.

I then proceed to discuss a 1932 study that examined the impact of the composition of cement on the heat released by the cement as it hardened (15).

Suppose the random variable Y represents not the heat from cement but the physiological thing you study: L-ascorbic acid transport, differential gene expression, TNF- α , or venous capacitance in trout. Assume that your Y is distributed normally with mean μ and standard deviation σ . We now have an example that is relevant to everyone. Unfortunately, we now also have a problem: different responses have different means and standard deviations. We can circumvent this problem if we consider the distribution of each response to be a standard normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$ (Fig. 2). This standard normal distribution is the population from which we will obtain our simulated sample observations—our data.

As the statistical cornerstone for our explorations, suppose we want to estimate $\mu = 0$ and $\sigma = 1$, the mean and standard deviation of our population (see Ref. 14). To do this, we draw at random a sample of n observations from the population. For simplicity, suppose we limit the sample to nine observations. This is the R command (*Advances_Statistics_Code.R*, line 36) that generates the sample and rounds each value to three decimal places:

```
TheData <- round ( rnorm ( nObs, mean = PopMean, sd = PopSD), 3)
```

The sample size is defined by the command `nObs <- 9` (*Advances_Statistics_Code.R*, line 10).

Because we had so much fun taking 1 random sample, we repeat the process until we have drawn a total of 1000 random samples, each with 9 observations, from our population. Mercifully, the command for `(i in 1:nSamples)` in line 35 of

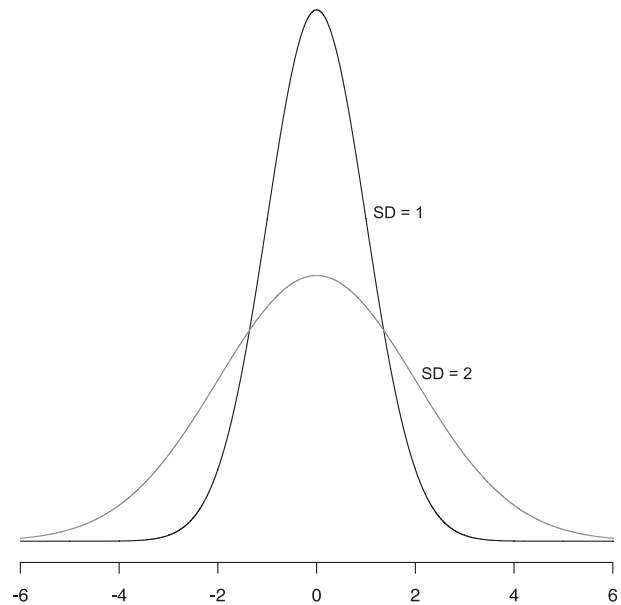


Fig. 4. Two normal distributions. These distributions differ in variability, characterized by the standard deviation SD. The commands in lines 109–129 of *Advances_Statistics_Code.R* create these distributions and this data graphic. To generate this data graphic, highlight and submit the lines of code from Figure 4: first line to Figure 4: last line.

Advances_Statistics_Code.R does this for us. These are the observations—the data—for samples 1, 2, and 1000:

```
> # Sample Observations
[1] 0.422 1.103 1.006 1.034 0.285 -0.647 1.235 0.912 1.825
[2] 0.154 -0.654 -0.147 1.715 0.720 0.804 0.256 1.155 0.646
:
[1000] 0.560 -1.138 0.485 -0.864 -0.277 2.198 0.050 0.500 0.587
```

Your sample observations will differ.

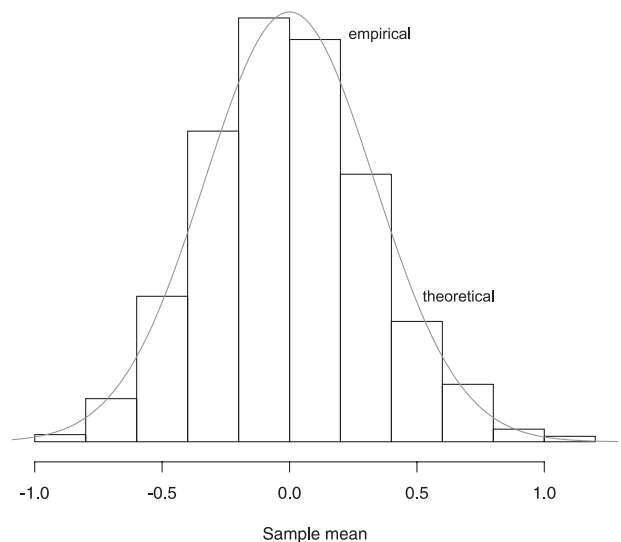


Fig. 5. Empirical (black) and theoretical (gray) distributions of the sample mean for 9 observations. The empirical distribution is composed of 1000 sample means. The empirical standard deviation of the 1000 sample means, 0.326, is near the theoretical value of 1/3. The commands in lines 138–139 of *Advances_Statistics_Code.R* calculate the theoretical distribution of the sample mean, and the commands in lines 141–150 create this data graphic. To generate this data graphic, highlight and submit the lines of code from Figure 5: first line to Figure 5: last line.

We have our data, but if we want to really understand the distinction between standard deviation and standard error, we also need some sample statistics.⁶ So each time we draw a sample of 9 observations, we calculate the sample statistics listed in Table 1. These are the statistics for *samples 1, 2, and 1000*:

> # Sample	Mean	SD	SE	t	LCI	UCI
[.1]	[.2]	[.3]	[.4]	[.5]	[.6]	[.7]
1	0.797	0.702	0.234	3.407	0.362	1.232
2	0.517	0.707	0.236	2.193	0.079	0.955
:						
1000	0.233	0.975	0.325	0.718	-0.371	0.838

The commands in *lines 35–62* of *Advances_Statistics_Code.R* compute these statistics.⁷

With these 1000 sets of sample observations and statistics, we are ready to explore the essential distinction between standard deviation and standard error.

Standard Deviation

In each of our 1000 samples, the 9 observations differ because the underlying population (see Fig. 2) is distributed over a range of possible values. The typical measure of the variability among experimental measurements is the sample standard deviation s :

$$s = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n - 1}}$$

where n is the number of observations in the sample, y_i is an individual observation, and \bar{y} is the sample mean. The sample standard deviation characterizes the dispersion of observations about the sample mean and estimates the population standard deviation σ . For example, the standard deviation of the observations in *sample 1*, 0.422, 1.103, . . . , 1.825, is $s = 0.702$, which estimates $\sigma = 1$. The empirical distribution of the 1000 sample standard deviations is centered at 0.966, slightly less than the actual value of 1 (Fig. 3). The command in *line 104* of *Advances_Statistics_Code.R* returns this value. Your value will differ slightly.

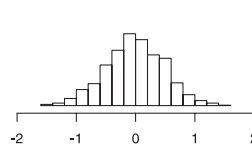
A larger standard deviation means greater dispersion: more variability (Fig. 4).

Standard Error of the Mean

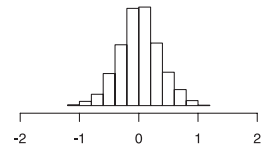
In words, what is the standard error of the mean $SE\{\bar{y}\}$? I ask this question of my students on the first day of class. Often students can explain in words how to calculate the standard error: divide the standard deviation by the square root of the sample size. Seldom can a student explain the concept behind the standard error: if I repeat an experiment a whole bunch of times—and each time I calculate a sample mean—then the standard deviation of those sample means will be the standard error of the mean (12–14). The standard error of the mean answers a theoretical question: if I repeat an experiment a whole bunch of times, by how much will a typical sample mean differ from the population mean?

By virtue of our simulation, we have 1000 sample means (Fig. 5). If we treat these 1000 sample means as observations, we can calculate their average and standard deviation:

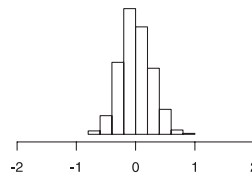
A. $n = 4$: $\text{Var}\{\text{means}\} = 0.24$



B. $n = 8$: $\text{Var}\{\text{means}\} = 0.12$



C. $n = 16$: $\text{Var}\{\text{means}\} = 0.07$



D. $n = 32$: $\text{Var}\{\text{means}\} = 0.03$

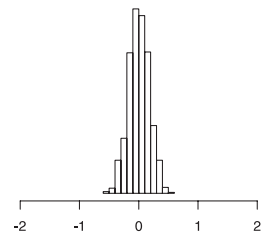


Fig. 6. Empirical distributions of sample means. These distributions are composed of 1000 samples of 4 (A), 8 (B), 16 (C), and 32 (D) observations drawn at random from a normal population for which the mean $\mu = 0$ and the standard deviation $\sigma = 1$ (see Fig. 1). For each distribution, the average of the sample means, $\text{Ave}\{\bar{y}\}$, is 0. As sample size increases, the sample means cluster more closely about $\text{Ave}\{\bar{y}\}$. Each time the sample size doubles, the variance of the sample means, $\text{Var}\{\text{means}\} \doteq \text{Var}\{\bar{y}\}$, is roughly halved. The commands in *lines 165–178* of *Advances_Statistics_Code.R* generate this simulation, and the commands in *lines 210–253* create this data graphic. To generate this data graphic, highlight and submit the lines of code from Figure 6: first line to Figure 6: last line.

$$\text{Ave}\{\bar{y}\} = 0.0 \quad \text{and} \quad \text{SD}\{\bar{y}\} = 0.326$$

The commands in *lines 156–157* of *Advances_Statistics_Code.R* return these values. Your values will differ slightly.

Suppose we draw from our population (see Fig. 2) an infinite number of samples, each with $n = 9$ observations. The infinite number of sample means, $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_\infty$, will be distributed normally with mean μ and standard deviation σ/\sqrt{n} .⁸ In other words, the average of the sample means, $\text{Ave}\{\bar{y}\}$, will be the population mean μ , but the standard deviation of the sample means, $\text{SD}\{\bar{y}\}$, will be smaller than the population standard deviation σ by a factor of σ/\sqrt{n} :

$$\text{Ave}\{\bar{y}\} = \mu \quad \text{and} \quad \text{SD}\{\bar{y}\} = \sigma/\sqrt{n} = 1/3$$

If the sample size n increases, then the standard deviation of the theoretical distribution of the sample mean will decrease: the more sample observations we have, the more certain we will be that the sample mean \bar{y} is near the actual population mean μ (Fig. 6).

⁸ I derive these results for the mean and standard deviation in Ref. 14. The Central Limit Theorem states that the theoretical distribution of the sample mean will be approximately normal regardless of the distribution of the original observations. If the distribution of the original observations happens to be normal, then the theoretical distribution of the sample mean will be exactly normal.

⁶ A statistic is a quantity calculated from the sample observations.

⁷ We will use the statistics in *columns 4–7* in subsequent explorations.

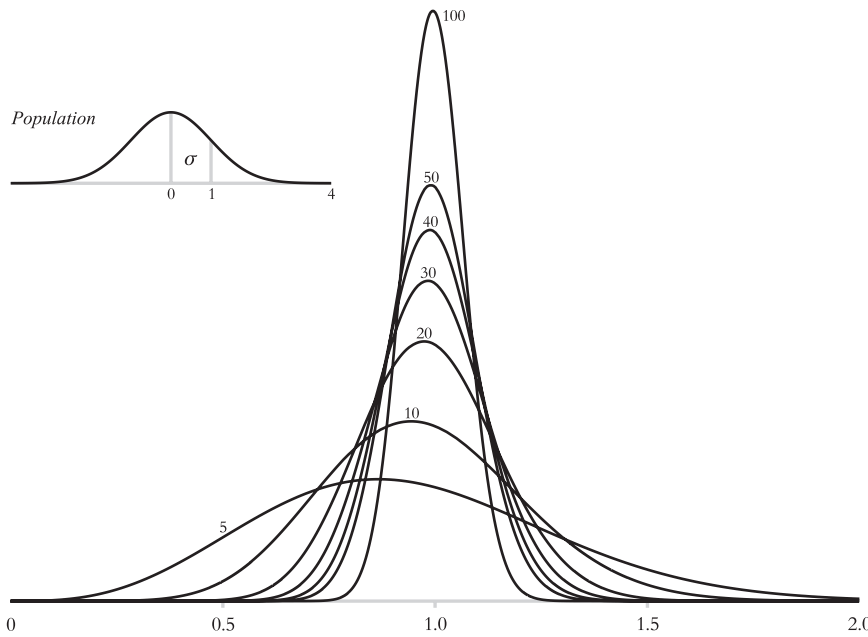


Fig. 7. The probability density function—the theoretical distribution of possible values—of the sample standard deviation for 5, 10, 20, 30, 40, 50, and 100 observations drawn from a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$ (inset). Equation A1 describes these probability density functions.

The standard deviation of the distribution of the sample mean is the standard error of the sample mean $SE\{\bar{y}\}$.

Summary

As this exploration has demonstrated, the standard deviation and standard error of the mean estimate quite different things: a standard deviation estimates the variability among individual observations in a sample—it can also estimate the variability in the underlying population—but a standard error of the mean estimates the theoretical variability among sample means.

In a sample, the observations—the data—differ because the population from which they were drawn is distributed over a range of possible values. The standard deviation describes the dispersion of these sample observations about the sample mean. If we fail to report the standard deviation, then we fail to fully report our data. Because it incorporates information about sample size, the standard error of the mean is a misguided estimate of variability among observations. Instead, the standard error of the mean provides an estimate of the uncertainty of the true value of the population mean.

In the next installment of this series, we will explore some of the concepts behind hypothesis testing: test statistics and P values.

APPENDIX

The probability density function—the theoretical distribution of possible values—of the sample standard deviation s is

$$f(s) = 2 \left(\frac{n^*}{\sigma^2} \right)^{n^*} \cdot \frac{s^{n^*-2}}{\Gamma(n^*)} \cdot e^{-n^* s^2/\sigma^2} \quad , \quad \text{for } 0 < s < \infty \quad , \quad (A1)$$

where $n^* = (n - 1)/2$ and the gamma-function $\Gamma(n^*)$ is

$$\Gamma(n^*) = \int_0^\infty w^{(n^*-1)} \cdot e^{-w} dw \quad .$$

If n^* is a positive integer greater than 1, then $\Gamma(n^*) = (n^* - 1)!$. For example, $\Gamma(4) = 3 \cdot 2 = 6$.

Figure 7 depicts the probability density function of the sample standard deviation for 5, 10, 20, 30, 40, 50, and 100 observations.

ACKNOWLEDGMENTS

I thank Matthew Strand (National Jewish Medical and Research Center, Denver, CO) for deriving the probability density function for the sample standard deviation.

REFERENCES

1. Altman DG. Analysing data. *Br Med J* 281: 1473–1475, 1980.
2. Altman DG. Collecting and screening data. *Br Med J* 281: 1399–1401, 1980.
3. Altman DG. How large a sample? *Br Med J* 281: 1336–1338, 1980.
4. Altman DG. Interpreting results. *Br Med J* 281: 1612–1614, 1980.
5. Altman DG. Misuse of statistics is unethical. *Br Med J* 281: 1182–1184, 1980.
6. Altman DG. Presentation of results. *Br Med J* 281: 1542–1544, 1980.
7. Altman DG. Study design. *Br Med J* 281: 1267–1269, 1980.
8. Altman DG, Gardner MJ. Presentation of variability. *Lancet* 2: 639, 1986.
9. American Physiological Society. *Instructions for Preparing Your Manuscript. Manuscript Sections* (online). http://www.the-aps.org/publications/i4a/prep_manuscript.htm#manuscript_sections [12 May 2008].
10. Brown GW. Standard deviation, standard error. Which “standard” should we use? *Am J Dis Child* 136: 937–941, 1982.
11. Curran-Everett D. Multiple comparisons: philosophies and illustrations. *Am J Physiol Regul Integr Comp Physiol* 279: R1–R8, 2000.
12. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society. *Am J Physiol Cell Physiol* 287: C243–C245, 2004.
13. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. *Adv Physiol Educ* 31: 295–298, 2007.
14. Curran-Everett D, Taylor S, Kafadar K. Fundamental concepts in statistics: elucidation and illustration. *J Appl Physiol* 85: 775–786, 1998.
15. Draper NR, Smith H. *Applied Regression Analysis*. New York: Wiley, 1981, p. 629.
16. Eisenhart C. Expression of the uncertainties of final results. *Science* 160: 1201–1204, 1968.
17. Gardner MJ. Understanding and presenting variation. *Lancet* 1: 230–231, 1975.
18. Healy MJR. Populations and samples. *Arch Dis Child* 66: 1355–1356, 1991.
19. Healy MJR. Significance tests. *Arch Dis Child* 66: 1457–1458, 1991.
20. Healy MJR. Data structures. *Arch Dis Child* 67: 533–535, 1992.

21. **Healy MJR.** Data structures (continued). *Arch Dis Child* 67: 757–759, 1992.
22. **Healy MJR.** Estimation. *Arch Dis Child* 67: 149–150, 1992.
23. **Healy MJR.** Regression and correlation. *Arch Dis Child* 67: 1306–1309, 1992.
24. **Healy MJR.** Sampling distributions. *Arch Dis Child* 67: 249–250, 1992.
25. **Healy MJR.** Size, power, and confidence. *Arch Dis Child* 67: 1495–1497, 1992.
26. **Healy MJR.** Counted data. *Arch Dis Child* 68: 246–249, 1993.
27. **Healy MJR.** Counted data (2). *Arch Dis Child* 68: 800–802, 1993.
28. **Healy MJR.** Data transformations. *Arch Dis Child* 69: 260–264, 1993.
29. **Healy MJR.** Non-normal data. *Arch Dis Child* 70: 158–163, 1994.
30. **Healy MJR.** Probability and decisions. *Arch Dis Child* 71: 90–94, 1994.
31. **Healy MJR.** Diagnostic and screening tests and reference values. *Arch Dis Child* 72: 358–361, 1995.
32. **Healy MJR.** Multiple regression (1). *Arch Dis Child* 73: 177–181, 1995.
33. **Healy MJR.** Multiple regression (2). *Arch Dis Child* 73: 270–274, 1995.
34. **R Development Core Team.** *R: a Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008; <http://www.R-project.org>.