Rounding numbers

Eugene H. Blackstone, MD

From the Department of Thoracic and Cardiovascular Surgery, Heart and Vascular Institute, and Department of Quantitative Health Sciences, Research Institute, Cleveland Clinic, Cleveland, Ohio.

Supported by the Department of Thoracic and Cardiovascular Surgery, Cleveland Clinic, Cleveland, Ohio.
Disclosures: Author has nothing to disclose with regard to commercial support.
Received for publication Sept 6, 2016; accepted for publication Sept 6, 2016; available ahead of print Oct 7, 2016.

Address for reprints: Eugene H. Blackstone, MD, Department of Thoracic and Cardiovascular Surgery, Cleveland Clinic, 9500 Euclid Ave, Desk JJ-4, Cleveland, OH 44195 (E-mail: blackse@ccf.org).

J Thorac Cardiovasc Surg 2016;152:1481-3
0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2016.09.003

The UNIVARIATE Procedure
Variable: hdl_pr (Preop HDL Cholesterol (mg/dl))

Moments

N	977	Sum Weights	977
Mean	39.7911975	Sum Observations	38876
Std Deviation	10.2347149	Variance	104.74939
Skewness	0.90179703	Kurtosis	1.90964934
Uncorrected SS	1649158	Corrected SS	102235.404
Coeff Variation	25.7210528	Std Error Mean	0.32743754

Sample printout of summary statistics.

Central Message

Expressing summary statistics and estimates from statistical analyses in scientific articles follows a logical set of little-known rules.

For many years I served on the examination committee for admission to the PhD program in statistics at the University of Alabama at Birmingham. I suppose for lack of creativity, committee members asked the same set of questions during each oral examination. One of my questions was this: "I see all these numbers with lots of decimal places on your statistics printout (Figure 1); how do you tell your investigators to express these numbers in their manuscripts?" The most common answer was that no one had ever taught them any rules for either summary statistics, like mean and standard deviation, or estimates from statistical analyses; they just gave the printout to the investigators. From the tabulated numbers encountered in the articles in the *Journal*, I daresay that you, as investigators and authors, are as clueless as these graduate students!

Thus, with some reorganization, I have extracted from Chapter 6 of the Kirklin/Barratt-Boyes book *Cardiac Surgery*[1] some guidance about rounding numbers. With this information you can be smarter than most statisticians, let alone a fifth grader! Certain generally agreed upon conventions for rounding numbers exist, although they are not easily found in print.[2,3]

TERMINOLOGY

A prerequisite for knowing how to round numbers is an understanding of the terminology used in expressing numbers. What is a digit, an even number, an odd number, decimal format versus decimal place, and particularly, what is a significant digit (see Table 1)? Pay particular attention to the fundamental difference between the decimal place and a significant digit: The numbers 0.00036, 36, and 36,000,000 all have 2 significant digits—3 and 6—but different decimal places.

Using the nomenclature and definitions presented in Table 1, we can now march through the steps in rounding numbers.

Step 1: Determine the Number of Significant Digits to Save

The number of significant digits to save is suggested by the precision of the measuring instrument for individual numbers and variability observed in a series of numbers. Precision of a measurement refers to its repeatability, whereas accuracy relates to how close the measurement is to the truth. A measurement can be quite precise, but if the measuring device is off calibration, the measurement could be quite inaccurate (a phenomenon often referred to as bias). One generally cannot go wrong by expressing numbers to the precision of the measurement tool.

In *The Journal of Thoracic and Cardiovascular Surgery*, authors are asked to present descriptive data (typically

The UNIVARIATE Procedure
Variable: hdl_pr (Preop HDL Cholesterol (mg/dl))

Moments

N	977	Sum Weights	977
Mean	39.7911975	Sum Observations	38876
Std Deviation	10.2347149	Variance	104.74939
Skewness	0.90179703	Kurtosis	1.90964934
Uncorrected SS	1649158	Corrected SS	102235.404
Coeff Variation	25.7210528	Std Error Mean	0.32743754

FIGURE 1. Raw statistical printout summarizing several aspects of the distribution of preoperative high-density lipoprotein (*HDL*) cholesterol in a study of 977 patients. Note the variable number of digits and significant figures.

TABLE 1. Expressing numbers

Digit	A digit is any of the 10 Arabic number symbols, 0 through 9. Digits are also called numbers, numerals, or integers.
Number	Although *number* and *digit* are synonyms, a number is more generally applied to a series of digits, separators (commas or decimal points), and other notations (see ''Scientific notation'' in this Table) that together represent a numeric quantity.
Even numbers	The Arabic numerals beginning with 0, 2, 4, 6, 8, and so on. These are numbers divisible by 2 without a remainder.
Odd numbers	The Arabic numerals beginning with 1, 3, 5, 7, 9, and so on. These are numbers divisible by 2 with a remainder of exactly 1.
Decimal format	Most numbers in scientific work are expressed in decimal format; that is, in a numerical system based on multiples of 10 as the fundamental unit, called the base of the numbering system. Other systems were prominent in antiquity, such as base 60 in Babylonian times, existing now only as the basis for clock time. Bases other than 10 are used in computer systems, such as base 2, 8, or 16. Others have been suggested as having better arithmetic properties; however, the fact that humans have 10 fingers has played a dominant role in popularizing the decimal system.
	In decimal notation, each place is a multiple of 10 and is named. A symbol known as the decimal point (a period is used in the United States) separates what is known as the units or ones place from the tenths (1/10) place. Whole numbers are to the left of the decimal point and fractional numbers to the right.
	To the left of the decimal point, the first place is called the units (or ones) place, the second the tens place, the third the hundreds place, and the fourth the thousands place. Commonly, a separator is inserted for each multiple of 1000 (a comma in the United States, and period in Europe; see below). In the number 1234, the 4 is in the units place, 3 in the tens place, 2 in the hundreds place, and 1 in the thousands place. Another way to express this number is as the sum of $4 + 30 + 200 + 1000$.
	To the right of the decimal point, the first place is called the tenths (1/10) place, the second the hundredths (1/100) place, and the third the thousandths (1/1000) place. Thus, in 0.1234, the 1 is in the tenths place, 2 in the hundredths place, 3 in the thousandths place, and 4 in the 10-thousandths place. Another way to express this number is as the sum of $0.1 + 0.02 + 0.003 + 0.0004$.
Decimal place	In the decimal system, decimal place is the position of digits immediately to the right of the symbol designating the decimal point. Location of the decimal point reflects the scale of measurement and is unrelated to significant digits.
Significant digit	Digits of the decimal form of a number beginning with the leftmost nonzero digit and extending to the right, with the implicit implication that all digits to the right are significant digits. That is, they are warranted either by inherent properties of the measuring device used to generate the numbers or by statistical properties of a collection of such numbers. Thus, in the number 23.56, the first significant digit is 2, the second is 3, the third is 5, and the fourth is 6. In the number 0.002356, the first significant digit is still 2, the second 3, the third 5, and the fourth 6. This emphasizes the fact that a significant digit is completely unrelated to the decimal place.
Leading zero	A zero placed before a decimal point is not considered a significant digit. It is generally used when it is implied that a nonzero significant digit could replace it, or to separate a negative sign ($-$), a positive sign ($+$), or a plus or minus sign ($\pm$) from the decimal point. Increasingly, numbers that are constrained to the range 0-1, such as probabilities (including *P* values), are expressed (displayed) without a leading zero.
Trailing zero	Trailing zeros are ambiguous. They may or may not be significant figures. For example, we have rounded \$72,347.23 to \$70,000. The zeros are just fillers up to the decimal point. To overcome this ambiguity, scientific notation was introduced.
Scientific notation	A method of expressing (displaying) numbers from 1-9, followed by a decimal point, the remaining significant digits, if any, multiplied by a power of 10. For example, 0.00037 in scientific notation is 3.7×10^{-4}, where 10^{-4} is 0.0001. In general, the numeric value, here 3.7, is called the mantissa, 10 is called the radix, and –4 is called the exponent. In our example of ambiguous zeros, the ambiguity in \$70,000 is eliminated by expressing it as $\$7 \times 10^4$.

presented as the article's Table 1) as mean ± standard deviation (if the distribution of data is roughly bell shaped—the so-called normal distribution). Standard deviation is an expression of the inherent variation of the numbers. In contrast, numbers such as regression coefficients are generally accompanied by their standard error, as the expression of the precision of those derived estimates. These measures of precision are the key to determining the number of significant digits to save.

For convenience, we will describe significant digits in terms of the mean and standard deviation. For cases in which the standard error is used, just translate ''standard deviation'' into ''standard error.''

Let's begin!

The place of the first significant digit of the standard deviation is found, and the mean or proportion is rounded to that place. The same place is saved in confidence limits. The standard deviation generally will be expressed to 1 additional place. Here are some examples:

A. Mean age is 72.17986, and the standard deviation is 9.364132. Nine is the first significant figure of the standard deviation and is in the ones place. Thus, we will keep 2 significant digits in the mean and 2 in the standard deviation: 72 ± 9.4.
B. Mean cost is \$72,347.23, and the standard deviation is \$23,994.06. The 2 in the 10-thousands place is the first significant digit in the standard deviation. Thus, the

mean cost would be expressed to 1 significant digit, and the standard deviation to 2 significant digits: \$70,000 ± \$24,000.

For Table 1 clinical data, the variation from patient to patient is generally such that mean and standard deviation are typically expressed to 2 significant figures.

Step 2: Look for Exceptions

Exceptions to Step 1 are as follows: (1) if the first significant digit of the standard deviation is 1, then 1 additional significant digit in the mean or proportion and standard deviation may be saved; (2) for percentages between 0% and 10% or between 90% and 100%, keep at least 2 significant digits; and (3) within a single table, consistency in saving digits may be desirable, so all numbers may be rounded to the place indicated by the majority of the numbers.

Step 3: Round the Numbers

Round the number by removing digits from its right side that falsely suggest a high degree of precision. This is done as follows[4]:

- If the digit in the first place beyond (to the right of) the significant digit to be rounded is $>$ 5, add 1 to the right-most digit to be retained and drop all other digits to its right. This is called rounding up. Thus, 2.77 to 2 significant figures would be rounded to 2.8. Similarly, 1479.336 to 2 significant figures would be rounded to 1500 and 0.000649376 would be rounded to 0.00065.
- If the digit in the first place beyond the significant digit to be rounded is $<$ 5, simply drop it and all other digits to its right. This is called rounding down. Thus, 3.44 to 2 significant figures would be rounded to 3.4. Similarly, 98,432.19 would be rounded to 98,000 and 0.00013175 would be rounded to 0.00013.
- If the digit in the first place beyond the digit to be rounded is exactly 5, add 1 to the rightmost digit to be retained if the last significant digit is odd (ie, 1, 3, 5, 7, or 9), and leave the digit to be rounded as is if it is even (ie, 0, 2, 4, 6, or 8). This rule results in the rightmost significant digit always being an even number.[5] Thus, 9.450000 to 2 significant figures would be rounded to 9.4, but 9.750000 would be rounded to 9.8.

CAVEAT

This discussion of rounding applies to summary statistics and derived quantities such as regression coefficients. It does not apply to the collection of data and the manipulation of those data (such as finding a mean of a series of numbers). One should retain all precision until the very end when one wishes to report the findings in an article.

References

1. Kouchoukos NT, Blackstone EH, Hanley FL, Kirklin JK. *Cardiac surgery*. 4th ed. Philadelphia: Elsevier; 2012.
2. Eisenhart C. Expression of the uncertainties of final results: clear statements of the uncertainties of reported values are needed for their critical evaluation. *Science*. 1968;160:1201-4.
3. Kirklin JW, Blackstone EH. Notes from the editors: tables. *J Thorac Cardiovasc Surg*. 1994;107:969.
4. Croxton FE. *Elementary statistics with applications in medicine and the biological sciences*. New York: Dover; 1953.
5. Midden WR. Rounding numbers: why the ''new system'' doesn't work [abstract]. *J Chem Educ*. 1997;74:405.