GLOBAL EDITION

Statistics for the Life Sciences

FIFTH EDITION

Myra L. Samuels
Jeffrey A. Witmer
Andrew A. Schaffner

ALWAYS LEARNING

PEARSON

Statistics for the Life Sciences

Fifth Edition

Global Edition

Myra L. Samuels

Purdue University

Jeffrey A. Witmer

Oberlin College

Andrew A. Schaffner

California Polytechnic State University,
San Luis Obispo

PEARSON

Boston Columbus Indianapolis New York San Francisco Hoboken
Amsterdam Cape Town Dubai London Madrid Milan Munich
Paris Montréal Toronto Delhi Mexico City São Paulo
Sydney Hong Kong Seoul Singapore Taipei Tokyo

ONE-SIDED CONFIDENCE INTERVALS

Most confidence intervals are of the form "estimate $\pm$ margin of error"; these are known as two-sided intervals. However, it is possible to construct a one-sided confidence interval, which is appropriate when only a lower bound, or only an upper bound, is of interest. The following two examples illustrate 90% and 95% one-sided confidence intervals.

Example 6.3.6 **Seeds per Fruit—One-Sided, 90%** Consider the seed data from Example 6.3.5, which are used to estimate the number of seeds per fruit for *Vallisneria americana.* It might be that we want a lower bound on μ, the population mean, but we are not concerned with how large μ might be. Whereas a two-sided 90% confidence interval is based on capturing the middle 90% of a t distribution and thus uses the t multipliers of $\pm t_{0.05}$, a one-sided 90% (lower) confidence interval uses the fact that $\Pr(-t_{0.10} < t < \infty) = 0.90$. Thus, the lower limit of the confidence interval is $\bar{y} - t_{0.10}\,\mathrm{SE}_{\bar{Y}}$ and the upper limit of the interval is infinity. In this case, with 11 degrees of freedom the t multiplier is $t_{11,\,0.10} = 1.363$ and we get

$$320 - 1.363(36) = 320 - 49 = 271$$

as the lower limit. The resulting interval is $(271, \infty)$. Thus, *we are 90% confident that the (population) mean number of seeds per fruit for Vallisneria americana is at least 271.* ■

Example 6.3.7 **Seeds per Fruit—One-Sided, 95%** A one-sided 95% confidence interval is constructed in the same manner as a one-sided 90% confidence interval, but with a different t multiplier. For the *Vallisneria americana* seeds data we have $t_{11,\,0.05} = 1.796$ and we get

$$320 - 1.796(36) = 320 - 65 = 255$$

as the lower limit. The resulting interval is $(255, \infty)$. Thus, *we are 95% confident that the (population) mean number of seeds per fruit for Vallisneria americana is at least 255.* ■

Exercises 6.3.1–6.3.22

6.3.1 **(Sampling exercise)** Refer to Exercise 5.2.1. Use your sample of five ellipse lengths to construct an 80% confidence interval for μ, using the formula $\bar{y} \pm (1.533)s/\sqrt{n}$.

6.3.2 **(Sampling exercise)** Refer to Exercise 5.2.3. Use your sample of 20 ellipse lengths to construct an 80% confidence interval for μ using the formula $\bar{y} \pm (1.328)s/\sqrt{n}$.

6.3.3 As part of a study of the development of the thymus gland, researchers weighed the glands of five chick embryos after 14 days of incubation. The thymus weights (mg) were as follows[12]:

29.6 21.5 28.0 34.6 44.9

For these data, the mean is 31.7 and the standard deviation is 8.7.

(a) Calculate the standard error of the mean.

(b) Construct a 90% confidence interval for the population mean.

6.3.4 Consider the data from Exercise 6.3.3.

(a) Construct a 95% confidence interval for the population mean.

(b) Interpret the confidence interval you found in part (a). That is, explain what the numbers in the interval mean. (See Examples 6.3.4 and 6.3.5.)

6.3.5 Six healthy 3-year-old female Suffolk sheep were injected with the antibiotic Gentamicin, at a dosage of 10 mg/kg body weight. Their blood serum concentrations (µg/ml) of Gentamicin 1.5 hours after injection were as follows[13]:

33 26 34 31 23 25

For these data, the mean is 28.7 and the standard deviation is 4.6.

(a) Construct a 95% confidence interval for the population mean.

(b) Define in words the population mean that you estimated in part (a). (See Example 6.1.1.)

(c) The interval constructed in part (a) nearly contains all of the observations; will this typically be true for a 95% confidence interval? Explain.

6.3.6 A zoologist measured tail length in 86 individuals, all in the 1-year age group, of the deermouse *Peromyscus*. The mean length was 60.43 mm and the standard deviation was 3.06 mm. A 95% confidence interval for the mean is (59.77, 61.09).

(a) True or false (and say why): We are 95% confident that the average tail length of the 86 individuals in the sample is between 59.77 mm and 61.09 mm.

(b) True or false (and say why): We are 95% confident that the average tail length of all the individuals in the population is between 59.77 mm and 61.09 mm.

6.3.7 Refer to Exercise 6.3.6.

(a) Without doing any computations, would an 80% confidence interval for the data in Exercise 6.3.6 be wider, narrower, or about the same? Explain.

(b) Without doing any computations, if 500 mice were sampled rather than 86, would the 95% confidence interval listed in Exercise 6.3.6 be wider, narrower, or about the same? Explain.

6.3.8 Researchers measured the bone mineral density of the spines of 94 women who had taken the drug CEE. (See Example 6.3.4, which dealt with hip bone mineral density.) The mean was 1.016 g/cm^2 and the standard deviation was 0.155 g/cm^2. A 95% confidence interval for the mean is (0.984, 1.048).

(a) True or false (and say why): 95% of the sampled bone mineral density measurements are between 0.984 and 1.048.

(b) True or false (and say why): 95% of the population bone mineral density measurements are between 0.984 and 1.048.

6.3.9 There was a control group in the study described in Example 6.3.4. The 124 women in the control group were given a placebo, rather than an active medication. At the end of the study they had an average bone mineral density of 0.840 g/cm^2. Shown are three confidence intervals: One is a 90% confidence interval, one is an 85% confidence interval, and the other is an 80% confidence interval. Without doing any calculations, match the intervals with the confidence levels and explain how you determined which interval goes with which level.

Confidence levels:

90% 85% 80%

Intervals (in scrambled order):

(0.826, 0.854) (0.824, 0.856) (0.822, 0.858)

6.3.10 Levels of insoluble ash (gm/kg) were measured in a sample of small (0.3 gm) aliquots of dried and ground alfalfa.[6] Below are three confidence intervals for the mean: One is a 95% confidence interval, one is a 90% confidence interval, and the other is an 85% confidence interval. Without doing any calculations, match the intervals with the confidence levels and explain how you determined which interval goes with which level.

Confidence levels:

95% 90% 85%

Intervals (in scrambled order):

(8.16, 10.88) (8.38, 10.66) (7.75, 11.29)

6.3.11 Human beta-endorphin (HBE) is a hormone secreted by the pituitary gland under conditions of stress. A researcher conducted a study to investigate whether a program of regular exercise might affect the resting (unstressed) concentration of HBE in the blood. He measured blood HBE levels, in January and again in May, from 10 participants in a physical fitness program. The results were as shown in the table.[14]

(a) Construct a 95% confidence interval for the population mean difference in HBE levels between January and May. (*Hint*: You need to use only the values in the right-hand column.)

	HBE Level (pg/ml)		
Participant	January	May	Difference
1	42	22	20
2	47	29	18
3	37	9	28
4	9	9	0
5	33	26	7
6	70	36	34
7	54	38	16
8	27	32	−5
9	41	33	8
10	18	14	4
Mean	37.8	24.8	13.0
SD	17.6	10.9	12.4

(b) Interpret the confidence interval from part (a). That is, explain what the interval tells you about HBE levels. (See Examples 6.3.4 and 6.3.5.)

(c) Using your interval to support your answer, is there evidence that HBE levels are lower in May than January? (*Hint*: Does your interval include the value zero?)

6.3.12 Consider the data from Exercise 6.3.11. If the sample size is small, as it is in this case, then in order for a confidence interval based on Student's t distribution to be valid, the data must come from a normally distributed population. Is it reasonable to think that difference in HBE level is normally distributed? How do you know?

6.3.13 Invertase is an enzyme that may aid in spore germination of the fungus *Colletotrichum graminicola.* A botanist incubated specimens of the fungal tissue in petri dishes and then assayed the tissue for invertase activity. The specific activity values for nine petri dishes incubated at 90% relative humidity for 24 hours are summarized as follows[15]:

$$\text{Mean} = 5{,}111 \text{ units} \quad \text{SD} = 818 \text{ units}$$

(a) Assume that the data are a random sample from a normal population. Construct a 95% confidence interval for the mean invertase activity under these experimental conditions.

(b) Interpret the confidence interval you found in part (a). That is, explain what the numbers in the interval mean. (See Examples 6.3.4 and 6.3.5.)

(c) If you had the raw data, how could you check the condition that the data are from a normal population?

6.3.14 As part of a study of the treatment of anemia in cattle, researchers measured the concentration of selenium in the blood of 36 cows who had been given a dietary supplement of selenium (2 mg/day) for 1 year. The cows were all the same breed (Santa Gertrudis) and had borne their first calf during the year. The mean selenium concentration was 6.21 μg/dl and the standard deviation was 1.84 μg/dl.[16] Construct a 95% confidence interval for the population mean.

6.3.15 In a study of larval development in the tufted apple budmoth *(Platynota idaeusalis),* an entomologist measured the head widths of 50 larvae. All 50 larvae had been reared under identical conditions and had moulted six times. The mean head width was 1.20 mm and the standard deviation was 0.14 mm. Construct a 90% confidence interval for the population mean.[17]

6.3.16 In a study of the effect of aluminum intake on the mental development of infants, a group of 92 infants who had been born prematurely were given a special aluminum-depleted intravenous-feeding solution.[18] At age 18 months the neurologic development of the infants was measured using the Bayley Mental Development Index. (The Bayley Mental Development Index is similar to an IQ score, with 100 being the average in the general population.) A 95% confidence interval for the mean is (93.8, 102.1).

(a) Interpret this interval. That is, what does the interval tell us about neurologic development in the population of prematurely born infants who receive intravenous-feeding solutions?

(b) Does this interval indicate that the mean IQ of the sampled population is below the general population average of 100?

6.3.17 In Exercise 6.3.16 a 95% confidence interval of (93.8, 102.1) was given based on a sample of $n = 92$ infants. Suppose, for sake of argument, that we could change the sample size and we would still get a sample mean of 97.95 and a sample SD of 20.04. How large would the sample size, n, need to be in order for the 95% confidence interval to exclude 100?

6.3.18 A group of 101 patients with end-stage renal disease were given the drug epoetin.[19] The mean hemoglobin level of the patients was 10.3 (g/dl), with an SD of 0.9. Construct a 95% confidence interval for the population mean.

6.3.19 In Table 4 we find that $t_{0.025} = 1.960$ when $\text{df} = \infty$. Show how this value can be verified using Table 3.

6.3.20 Use Table 3 to find the value of $t_{0.0025}$ when $\text{df} = \infty$. (Do not attempt to interpolate in Table 4.)

6.3.21 Data are often summarized in this format: $\bar{y} \pm \text{SE}$. Suppose this interval is interpreted as a confidence interval. If the sample size is large, what would be the confidence level of such an interval? That is, what is the chance that an interval computed as

$$\bar{y} \pm (1.00)\text{SE}$$

will actually contain the population mean? [*Hint*: Recall that the confidence level of the interval $\bar{y} \pm (1.96)\text{SE}$ is 95%.]

6.3.22 (Continuation of Exercise 6.3.21)

(a) If the sample size is small but the population distribution is normal, is the confidence level of the interval $\bar{y} \pm \text{SE}$ larger or smaller than the answer to Exercise 6.3.21? Explain.

(b) How is the answer to Exercise 6.3.21 affected if the population distribution of Y is not approximately normal?

6.4 Planning a Study to Estimate μ

Before collecting data for a research study, it is wise to consider in advance whether the estimates generated from the data will be sufficiently precise. It can be painful indeed to discover after a long and expensive study that the standard errors are so large that the primary questions addressed by the study cannot be answered.

The precision with which a population mean can be estimated is determined by two factors: (1) the population variability of the observed variable Y, and (2) the sample size.

In some situations the variability of Y cannot, and perhaps should not, be reduced. For example, a wildlife ecologist may wish to conduct a field study of a natural population of fish; the heterogeneity of the population is not controllable and in fact is a proper subject of investigation. As another example, in a medical investigation, in addition to knowing the average response to a treatment, it may also be important to know how much the response varies from one patient to another, and so it may not be appropriate to use an overly homogeneous group of patients.

On the other hand, it is often appropriate, especially in comparative studies, to reduce the variability of Y by holding *extraneous* conditions as constant as possible. For example, physiological measurements may be taken at a fixed time of day; tissue may be held at a controlled temperature; all animals used in an experiment may be the same age.

Suppose, then, that plans have been made to reduce the variability of Y as much as possible, or desirable. What sample size will be sufficient to achieve a desired degree of precision in estimation of the population mean? If we use the standard error as our measure of precision, then this question can be approached in a straightforward manner. Recall that the SE is defined as

$$\mathrm{SE}_{\bar{Y}} = \frac{s}{\sqrt{n}}$$

In order to decide on a value of n, one must (1) specify what value of the SE is considered desirable to achieve and (2) have available a preliminary guess of the SD, either from a pilot study or other previous experience, or from the scientific literature. The required sample size is then determined from the following equation:

$$\text{Desired SE} = \frac{\text{Guessed SD}}{\sqrt{n}}$$

The following example illustrates the use of this equation.

Example 6.4.1 **Butterfly Wings** The butterfly wing data of Example 6.1.1 yielded the following summary statistics:

$$\bar{y} = 32.81 \text{ cm}^2$$
$$s = 2.48 \text{ cm}^2$$
$$\mathrm{SE} = 0.66 \text{ cm}^2$$

Suppose the researcher is now planning a new study of butterflies and has decided that it would be desirable that the SE be no more than 0.4 cm^2. As a preliminary guess of the SD, she will use the value from the old study, namely 2.48 cm^2. Thus, the desired n must satisfy the following relation:

$$\mathrm{SE} = \frac{2.48}{\sqrt{n}} \leq 0.4$$

This equation is easily solved to give $n \geq 38.4$. Since one cannot have 38.4 butterflies, the new study should include at least 39 butterflies. ■

You may wonder how a researcher would arrive at a value such as 0.4 cm^2 for the desired SE. Such a value is determined by considering how much error one is willing to tolerate in the estimate of μ. For example, suppose the researcher in Example 6.4.1 has decided that she would like to be able to estimate the population mean, μ, to within ± 0.8 with 95% confidence. That is, she would like her 95% confidence interval for μ to be $\bar{y} \pm 0.8$. The "$\pm$ part" of the confidence interval, which is sometimes called the **margin of error for 95% confidence,** is $t_{0.025} \times$ SE. The precise value of $t_{0.025}$ depends on the degrees of freedom, but typically $t_{0.025}$ is approximately 2. Thus, the researcher wants $2 \times$ SE to be no more than 0.8. This means that the SE should be no more than 0.4 cm^2.

In comparative studies, the primary consideration is usually the size of anticipated treatment effects. For instance, if one is planning to compare two experimental groups or distinct populations, the anticipated SE for each population or experimental group should be substantially smaller than (preferably less than one-fourth of) the anticipated difference between the two group means.* Thus, the butterfly researcher of Example 6.4.1 might arrive at the value 0.4 cm^2 if she were planning to compare male and female Monarch butterflies and she expected the wing areas for the sexes to differ (on the average) by about 1.6 cm^2. She would then plan to capture 39 male and 39 female butterflies.

To see how the required n depends on the specified precision, suppose the butterfly researcher specified the desired SE to be 0.2 cm^2 rather than 0.4 cm^2. Then the relation would be

$$\text{SE} = \frac{2.48}{\sqrt{n}} \leq 0.2$$

which yields $n = 153.76$, so she would plan to capture 154 butterflies of each sex. Thus, to double the precision (by cutting the SE in half) requires not twice as many but four times as many observations. This phenomenon of "diminishing returns" is due to the square root in the SE formula.

*This is a rough guideline for obtaining adequate sensitivity to discriminate between treatments. Such sensitivity, technically called *power,* is discussed in Chapter 7.

Exercises 6.4.1–6.4.6

6.4.1 An experiment is being planned to compare the effects of several diets on the weight gain of beef cattle, mea-sured over a 140-day test period.[20] In order to have enough precision to compare the diets, it is desired that the standard error of the mean for each diet should not exceed 5 kg.

(a) If the population standard deviation of weight gain is guessed to be about 20 kg on any of the diets, how many cattle should be put on each diet in order to achieve a sufficiently small standard error?

(b) If the guess of the standard deviation is doubled, to 40 kg, does the required number of cattle double? Explain.

6.4.2 A medical researcher proposes to estimate the mean serum cholesterol level of a certain population of middle-aged men, based on a random sample of the population. He asks a statistician for advice. The ensuing discussion reveals that the researcher wants to estimate the population mean to within ± 6 mg/dl or less, with 95% confidence. Thus, the standard error of the mean should be 3 mg/dl or less. Also, the researcher believes that the standard deviation of serum cholesterol in the population is probably about 40 mg/dl.[21] How large a sample does the researcher need to take?

6.4.3 A plant physiologist is planning to measure the stem lengths of soybean plants after 2 weeks of growth when using a new fertilizer. Previous experiments suggest that the standard deviation of stem length is around 1.2 cm.[22] Using this as a guess of σ, determine how many soybean plants the researcher should have if she wants the standard error of the group mean to be no more than 0.2 cm.

6.4.4 Suppose you are planning an experiment to test the effects of various diets on the weight gain of young turkeys. The observed variable will be Y = weight gain in 3 weeks (measured over a period starting 1 week after hatching and ending 3 weeks later). Previous experiments suggest that the standard deviation of Y under a standard diet is approximately 80 g.[23] Using this as a guess of σ, determine how many turkeys you should have in a treatment group, if you want the standard error of the group mean to be no more than

(a) 20 g

(b) 15 g

6.4.5 A study of 29 female Sumatran elephants provided a 95% confidence interval for the mean shoulder height as (197.2, 235.1) cm.[24] Consider a new study to estimate the mean shoulder height of male Sumatran elephants.

Assuming the standard deviations of shoulder heights are similar for males and females, how many male elephants should be sampled so that the 95% confidence interval for the mean shoulder height of males will have a margin of error of 10 cm?

6.4.6 A researcher is planning to compare the effects of two different types of lights on the growth of bean plants. She expects that the means of the two groups will differ by about 1 inch and that in each group the standard deviation of plant growth will be around 1.5 inches. Consider the guideline that the anticipated SE for each experimental group should be no more than one-fourth of the anticipated difference between the two group means. How large should the sample be (for each group) in order to meet this guideline?

6.5 Conditions for Validity of Estimation Methods

For any sample of quantitative data, one can use the methods of this chapter to compute the mean, its standard error, and various confidence intervals; indeed, computers can make this rather easy to carry out. However, the *interpretations* that we have given for these descriptions of the data are valid only under certain conditions.

CONDITIONS FOR VALIDITY OF THE SE FORMULA

First, the very notion of regarding the sample mean as an estimate of a population mean requires that the data be viewed "as if" they had been generated by random sampling from some population. To the extent that this is not possible, any inference beyond the actual data is questionable. The following example illustrates the difficulty.

Example 6.5.1 **Marijuana and Intelligence** Ten people who used marijuana heavily were found to be quite intelligent; their mean IQ was 128.4, whereas the mean IQ for the general population is known to be 100. The 10 people belonged to a religious group that uses marijuana for ritual purposes. Since their decision to join the group might very well be related to their intelligence, it is not clear that the 10 can be regarded (with respect to IQ) as a random sample from any particular population, and therefore there is no apparent basis for thinking of the sample mean (128.4) as an estimate of the mean IQ of a particular population (e.g., all heavy marijuana users). An inference about the *effect* of marijuana on IQ would be even more implausible, especially because data were not available on the IQs of the 10 people *before* they began marijuana use.[25] ■

Second, the use of the standard error formula $\text{SE} = s/\sqrt{n}$ requires two further conditions:

1. The population size must be large compared to the sample size. This requirement is rarely a problem in the life sciences; the sample can be as much as 5% of the population without seriously invalidating the SE formula.*
2. The observations must be independent of each other. This requirement means that the n observations actually give n independent pieces of information about the population.

Data often fail to meet the independence requirement if the experiment or sampling regime has a **hierarchical structure,** in which observational units are "nested" within sampling units, as illustrated by the following example.

*If the sample size, n, is a substantial fraction of the population size, N, then the "finite population correction factor" should be applied. This factor is $\sqrt{\frac{N-n}{N-1}}$. The standard error of the mean then becomes $\frac{s}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$.