8

© Steve Rosset/iStockphoto

Statistical Intervals for a Single Sample

INTRODUCTION

Engineers are often involved in estimating parameters. For example, there is an ASTM Standard E23 that defines a technique called the Charpy V-notch method for notched bar impact testing of metallic materials. The impact energy is often used to determine if the material experiences a ductile-to-brittle transition as the temperature decreases. Suppose that you have tested a sample of 10 specimens of a particular material with this procedure. You know that you can use the sample average $\overline{X}$ to estimate the true mean impact energy μ. However, we also know that the true mean impact energy is unlikely to be exactly equal to your estimate. Reporting the results of your test as a single number is unappealing, because there is nothing inherent in $\overline{X}$ that provides any information about how close it is to μ. Your estimate could be very close, or it could be considerably far from the true mean. A way to avoid this is to report the estimate in terms of a range of plausible values called a confidence interval. A confidence interval always specifies a confidence level, usually 90%, 95%, or 99%, which is a measure of the reliability of the procedure. So if a 95% confidence interval on the impact energy based on the data from your 10 specimens has a lower limit of 63.84J and an upper limit of 65.08J, then we can say that at the 95% level of confidence any value of mean impact energy between 63.84 J and 65.08 J is a plausible value. By reliability, we mean that if we repeated this experiment over and over again, 95% of all samples would produce a confidence interval that contains the true mean impact energy, and only 5% of the time would the interval be in error. In this chapter you will learn how to construct confidence intervals and other useful types of statistical intervals for many important types of problem situations.

CHAPTER OUTLINE

LEARNING OBJECTIVES

After careful study of this chapter, you should be able to do the following:

1. Construct confidence intervals on the mean of a normal distribution, using either the normal distribution or the *t* distribution method
2. Construct confidence intervals on the variance and standard deviation of a normal distribution
3. Construct confidence intervals on a population proportion
4. Use a general method for constructing an approximate confidence interval on a parameter
5. Construct prediction intervals for a future observation
6. Construct a tolerance interval for a normal population
7. Explain the three types of interval estimates: confidence intervals, prediction intervals, and tolerance intervals

In the previous chapter we illustrated how a point estimate of a parameter can be estimated from sample data. However, it is important to understand how good is the estimate obtained. For example, suppose that we estimate the mean viscosity of a chemical product to be $\hat{\mu} = \bar{x} = 1000$. Now because of sampling variability, it is almost never the case that the true mean μ is exactly equal to the estimate $\bar{x}$. The point estimate says nothing about how close $\hat{\mu}$ is to μ. Is the process mean likely to be between 900 and 1100? Or is it likely to be between 990 and 1010? The answer to these questions affects our decisions regarding this process. Bounds that represent an interval of plausible values for a parameter are an example of an interval estimate. Surprisingly, it is easy to determine such intervals in many cases, and the same data that provided the point estimate are typically used.

An interval estimate for a population parameter is called a **confidence interval.** Information about the precision of estimation is conveyed by the length of the interval. A short interval implies precise estimation. We cannot be certain that the interval contains the true, unknown population parameter—we only use a sample from the full population to compute the point estimate and the interval. However, the confidence interval is constructed so that we have high confidence that it does contain the unknown population parameter. Confidence intervals are widely used in engineering and the sciences.

A **tolerance interval** is another important type of interval estimate. For example, the chemical product viscosity data might be assumed to be normally distributed. We might like to calculate limits that bound 95% of the viscosity values. For a normal distribution, we know that 95% of the distribution is in the interval

$$\mu - 1.96\sigma, \mu + 1.96\sigma$$

However, this is not a useful tolerance interval because the parameters μ and σ are unknown. Point estimates such as $\bar{x}$ and s can be used in the above equation for μ and σ. However, we need to account for the potential error in each point estimate to form a tolerance interval for the distribution. The result is an interval of the form

$$\bar{x} - ks, \bar{x} + ks$$

where k is an appropriate constant (that is larger than 1.96 to account for the estimation error). As in the case of a confidence interval, it is not certain that the tolerance interval bounds 95% of the distribution, but the interval is constructed so that we have high confidence that it does. Tolerance intervals are widely used and, as we will subsequently see, they are easy to calculate for normal distributions.

Confidence and tolerance intervals bound unknown elements of a distribution. In this chapter you will learn to appreciate the value of these intervals. A **prediction interval** provides bounds on one (or more) **future observations** from the population. For example, a prediction interval could be used to bound a single, new measurement of viscosity—another useful interval. With a large sample size, the prediction interval for normally distributed data tends to the tolerance interval, but for more modest sample sizes the prediction and tolerance intervals are different.

Keep the purpose of the three types of interval estimates clear:

- A confidence interval bounds population or distribution parameters (such as the mean viscosity).
- A tolerance interval bounds a selected proportion of a distribution.
- A prediction interval bounds future observations from the population or distribution.

8-1 CONFIDENCE INTERVAL ON THE MEAN OF A NORMAL DISTRIBUTION, VARIANCE KNOWN

The basic ideas of a confidence interval (CI) are most easily understood by initially considering a simple situation. Suppose that we have a normal population with unknown mean μ and known variance σ^2. This is a somewhat unrealistic scenario because typically both the mean and variance are unknown. However, in subsequent sections we will present confidence intervals for more general situations.

8-1.1 Development of the Confidence Interval and Its Basic Properties

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution with unknown mean μ and known variance σ^2. From the results of Chapter 5 we know that the sample mean

$\overline{X}$ is normally distributed with mean μ and variance σ^2/n. We may **standardize** $\overline{X}$ by subtracting the mean and dividing by the standard deviation, which results in the variable

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \tag{8-1}$$

The random variable Z has a standard normal distribution.

A **confidence interval** estimate for μ is an interval of the form $l \leq \mu \leq u$, where the end-points l and u are computed from the sample data. Because different samples will produce different values of l and u, these end-points are values of random variables L and U, respectively. Suppose that we can determine values of L and U such that the following probability statement is true:

$$P\{L \leq \mu \leq U\} = 1 - \alpha \tag{8-2}$$

where $0 \leq \alpha \leq 1$. There is a probability of $1 - \alpha$ of selecting a sample for which the CI will contain the true value of μ. Once we have selected the sample, so that $X_1 = x_1$, $X_2 = x_2, \ldots,$ $X_n = x_n$, and computed l and u, the resulting **confidence interval** for μ is

$$l \leq \mu \leq u \tag{8-3}$$

The end-points or bounds l and u are called the **lower-** and **upper-confidence limits,** respectively, and $1 - \alpha$ is called the **confidence coefficient.**

In our problem situation, because $Z = (\overline{X} - \mu)/(\sigma/\sqrt{n})$ has a standard normal distribution, we may write

$$P\left\{-z_{\alpha/2} \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right\} = 1 - \alpha$$

Now manipulate the quantities inside the brackets by (1) multiplying through by $\sigma/\sqrt{n}$, (2) subtracting $\overline{X}$ from each term, and (3) multiplying through by -1. This results in

$$P\left\{\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right\} = 1 - \alpha \tag{8-4}$$

From consideration of Equation 8-4, the lower and upper limits of the inequalities in Equation 8-4 are the lower- and upper-confidence limits L and U, respectively. This leads to the following definition.

Confidence Interval on the Mean, Variance Known

If $\bar{x}$ is the sample mean of a random sample of size n from a normal population with known variance σ^2, a $100(1 - \alpha)\%$ CI on μ is given by

$$\bar{x} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \bar{x} + z_{\alpha/2}\sigma/\sqrt{n} \tag{8-5}$$

where $z_{\alpha/2}$ is the upper $100\alpha/2$ percentage point of the standard normal distribution.

EXAMPLE 8-1 Metallic Material Transition

ASTM Standard E23 defines standard test methods for notched bar impact testing of metallic materials. The Charpy V-notch (CVN) technique measures impact energy and is often used to determine whether or not a material experiences a ductile-to-brittle transition with decreasing temperature. Ten measurements of impact energy (J) on specimens of A238 steel cut at 60ºC are as follows: 64.1, 64.7, 64.5, 64.6, 64.5, 64.3, 64.6, 64.8, 64.2, and 64.3. Assume that impact energy is normally distributed with $\sigma = 1J$. We want to find a 95% CI for μ, the mean impact energy. The required quantities are $z_{\alpha/2} = z_{0.025} = 1.96$, $n = 10$, $\sigma = 1$, and $\bar{x} = 64.46$. The resulting 95% CI is found from Equation 8-5 as follows:

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$64.46 - 1.96\frac{1}{\sqrt{10}} \le \mu \le 64.46 + 1.96\frac{1}{\sqrt{10}}$$

$$63.84 \le \mu \le 65.08$$

Practical Interpretation: Based on the sample data, a range of highly plausible values for mean impact energy for A238 steel at 60°C is $63.84J \le \mu \le 65.08J$.

Interpreting a Confidence Interval

How does one interpret a confidence interval? In the impact energy estimation problem in Example 8-1, the 95% CI is $63.84 \le \mu \le 65.08$, so it is tempting to conclude that μ is within this interval with probability 0.95. However, with a little reflection, it's easy to see that this cannot be correct; the true value of μ is unknown and the statement $63.84 \le \mu \le 65.08$ is either correct (true with probability 1) or incorrect (false with probability 1). The correct interpretation lies in the realization that a CI is a *random interval* because in the probability statement defining the end-points of the interval (Equation 8-2), L and U are random variables. Consequently, the correct interpretation of a $100(1 - \alpha)\%$ CI depends on the relative frequency view of probability. Specifically, if an infinite number of random samples are collected and a $100(1 - \alpha)\%$ confidence interval for μ is computed from each sample, $100(1 - \alpha)\%$ of these intervals will contain the true value of μ.

The situation is illustrated in Fig. 8-1, which shows several $100(1 - \alpha)\%$ confidence intervals for the mean μ of a normal distribution. The dots at the center of the intervals indicate the point estimate of μ (that is, $\bar{x}$). Notice that one of the intervals fails to contain the true value of μ. If this were a 95% confidence interval, in the long run only 5% of the intervals would fail to contain μ.

Now in practice, we obtain only one random sample and calculate one confidence interval. Since this interval either will or will not contain the true value of μ, it is not reasonable to attach a probability level to this specific event. The appropriate statement is that the observed interval $[l, u]$ brackets the true value of μ with **confidence** $100(1 - \alpha)$. This statement has a frequency interpretation; that is, we don't know if the statement is true for this specific sample, but the *method* used to obtain the interval $[l, u]$ yields correct statements $100(1 - \alpha)\%$ of the time.

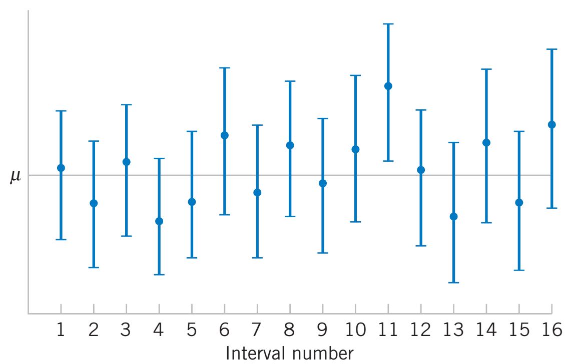

Figure 8-1 Repeated construction of a confidence interval for μ.

Confidence Level and Precision of Estimation

Notice in Example 8-1 that our choice of the 95% level of confidence was essentially arbitrary. What would have happened if we had chosen a higher level of confidence, say, 99%? In fact, doesn't it seem reasonable that we would want the higher level of confidence? At $\alpha = 0.01$, we find $z_{\alpha/2} = z_{0.01/2} = z_{0.005} = 2.58$, while for $\alpha = 0.05$, $z_{0.025} = 1.96$. Thus, the **length** of the 95% confidence interval is

$$2(1.96\sigma/\sqrt{n}) = 3.92\sigma/\sqrt{n}$$

whereas the length of the 99% CI is

$$2(2.58\sigma/\sqrt{n}) = 5.16\sigma/\sqrt{n}$$

Thus, the 99% CI is longer than the 95% CI. This is why we have a higher level of confidence in the 99% confidence interval. Generally, for a fixed sample size n and standard deviation σ, the higher the confidence level, the longer the resulting CI.

The length of a confidence interval is a measure of the **precision** of estimation. From the preceeding discussion, we see that precision is inversely related to the confidence level. It is desirable to obtain a confidence interval that is short enough for decision-making purposes and that also has adequate confidence. One way to achieve this is by choosing the sample size n to be large enough to give a CI of specified length or precision with prescribed confidence.

8-1.2 Choice of Sample Size

The precision of the confidence interval in Equation 8-5 is $2z_{\alpha/2}\sigma/\sqrt{n}$. This means that in using $\bar{x}$ to estimate μ, the error $E = |\bar{x} - \mu|$ is less than or equal to $z_{\alpha/2}\sigma/\sqrt{n}$ with confidence $100(1 - \alpha)$. This is shown graphically in Fig. 8-2. In situations where the sample size can be controlled, we can choose n so that we are $100(1 - \alpha)$ percent confident that the error in estimating μ is less than a specified bound on the error E. The appropriate sample size is found by choosing n such that $z_{\alpha/2}\sigma/\sqrt{n} = E$. Solving this equation gives the following formula for n.

Sample Size for Specified Error on the Mean, Variance Known

If $\bar{x}$ is used as an estimate of μ, we can be $100(1 - \alpha)\%$ confident that the error $|\bar{x} - \mu|$ will not exceed a specified amount E when the sample size is

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 \qquad (8\text{-}6)$$

If the right-hand side of Equation 8-6 is not an integer, it must be rounded up. This will ensure that the level of confidence does not fall below $100(1 - \alpha)\%$. Notice that $2E$ is the length of the resulting confidence interval.

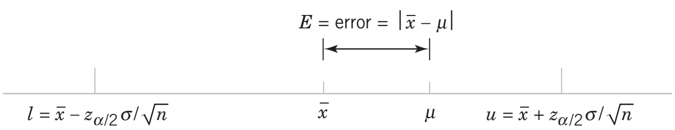

Figure 8-2 Error in estimating μ with $\bar{x}$.