8 Statistical Intervals for a Single Sample

CHAPTER OUTLINE

Learning Objectives for Chapter 8

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.

8-1 Introduction

- In the previous chapter we illustrated how a parameter can be estimated from sample data. However, it is important to understand how good is the estimate obtained.
- Bounds that represent an interval of plausible values for a parameter are an example of an interval estimate.
- Three types of intervals will be presented:
  - Confidence intervals
  - Prediction intervals
  - Tolerance intervals

8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known

8-2.1 Development of the Confidence Interval and its Basic Properties (Eq. 8-1)

$\overline{X}$ 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

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad (8\text{-}1)$$

The random variable Z has a standard normal distribution.

8-2.1 Development of the Confidence Interval and its Basic Properties (Eq. 8-2 & 3)

A **confidence interval** estimate for μ is an interval of the form $l \le \mu \le u$, 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:

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

where $0 \le \alpha \le 1$. 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

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

8-2.1 Development of the Confidence Interval and its Basic Properties (Eq. 8-4)

- The endpoints or bounds *l* and *u* are called lower- and upper-confidence limits, respectively.
- Since Z follows a standard normal distribution, we can write:

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

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

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

8-2.1 Development of the Confidence Interval and its Basic Properties (Eq. 8-5)

Definition

If $\overline{x}$ 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

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

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

Example 8-1

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 $\overline{x} = 64.46$. The resulting 95% CI is found from Equation 8-5 as follows:

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

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

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

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

- The confidence interval is a random interval
- The appropriate interpretation of a confidence interval (for example on μ) is: The observed interval $[l, u]$ brackets the true value of μ, with confidence $100(1-\alpha)$.
- Examine Figure 8-1 on the next slide.

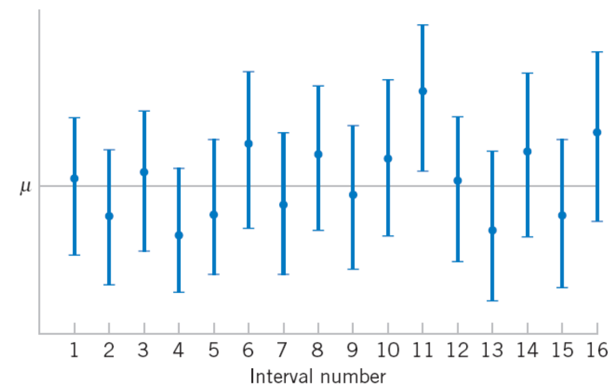

Figure 8-1 Repeated construction of a confidence interval for μ.

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

Confidence Level and Precision of Error

The length of a confidence interval is a measure of the precision of estimation.

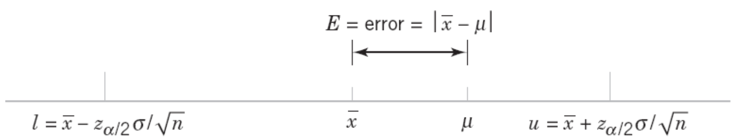

Figure 8-2 Error in estimating μ with $\bar{x}$.

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

8-2.2 Choice of Sample Size

(Eq. 8-6)

Definition

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

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

Example 8-2

EXAMPLE 8-2 Metallic Material Transition

To illustrate the use of this procedure, consider the CVN test described in Example 8-1, and suppose that we wanted to determine how many specimens must be tested to ensure that the 95% CI on μ for A238 steel cut at 60°C has a length of at most 1.0J. Since the bound on error in estimation E is one-half of the length of the CI, to determine n we use Equation 8-6 with

$E = 0.5$, $\sigma = 1$, and $z_{\alpha/2} = 1.96$. The required sample size is 16,

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 = \left[\frac{(1.96)1}{0.5}\right]^2 = 15.37$$

and because n must be an integer, the required sample size is $n = 16$.

8-2.3 One-Sided Confidence Bounds (Eq. 8-7 & 8)

Definition

A $100(1-\alpha)\%$ **upper-confidence bound** for μ is

$$\mu \le u = \bar{x} + z_\alpha \sigma/\sqrt{n} \qquad (8\text{-}7)$$

and a $100(1-\alpha)\%$ **lower-confidence bound** for μ is

$$\bar{x} - z_\alpha \sigma/\sqrt{n} = l \le \mu \qquad (8\text{-}8)$$

8-2.4 General Method to Derive a Confidence Interval

It is easy to give a general method for finding a confidence interval for an unknown parameter θ. Let $X_1, X_2, \ldots, X_n$ be a random sample of n observations. Suppose we can find a statistic $g(X_1, X_2, \ldots, X_n; \theta)$ with the following properties:

1. $g(X_1, X_2, \ldots, X_n; \theta)$ depends on both the sample and θ.
2. The probability distribution of $g(X_1, X_2, \ldots, X_n; \theta)$ does not depend on θ or any other unknown parameter.

8-2.4 General Method to Derive a Confidence Interval (Eq. 8-9 & 10)

In the case considered in this section, the parameter $\theta = \mu$. The random variable $g(X_1, X_2, \ldots, X_n; \mu) = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ and satisfies both conditions above; it depends on the sample and on μ, and it has a standard normal distribution since σ is known. Now one must find constants C_L and C_U so that

$$P[C_L \le g(X_1, X_2, \ldots, X_n; \theta) \le C_U] = 1 - \alpha \qquad (8\text{-}9)$$

Because of property 2, C_L and C_U do not depend on θ. In our example, $C_L = -z_{\alpha/2}$ and $C_U = z_{\alpha/2}$. Finally, you must manipulate the inequalities in the probability statement so that

$$P[L(X_1, X_2, \ldots, X_n) \le \theta \le U(X_1, X_2, \ldots, X_n)] = 1 - \alpha \qquad (8\text{-}10)$$

8-2.4 General Method to Derive a Confidence Interval

This gives $L(X_1, X_2, \ldots, X_n)$ and $U(X_1, X_2, \ldots, X_n)$ as the lower and upper confidence limits defining the $100(1-\alpha)\%$ confidence interval for θ. The quantity $g(X_1, X_2, \ldots, X_n; \theta)$ is often called a "pivotal quantity" because we pivot on this quantity in Equation 8-11 to produce Equation 8-12. In our example, we manipulated the pivotal quantity $(\overline{X}-\mu)/(\sigma/\sqrt{n})$ to obtain $L(X_1, X_2, \ldots, X_n) = \overline{X} - z_{\alpha/2}\sigma/\sqrt{n}$ and $U(X_1, X_2, \ldots, X_n) = \overline{X} + z_{\alpha/2}\sigma/\sqrt{n}$.

8-2.5 A Large-Sample Confidence Interval for μ (Eq. 8-11)

Definition

When n is large, the quantity

$$\frac{\overline{X}-\mu}{S/\sqrt{n}}$$

has an approximate standard normal distribution. Consequently,

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}} \qquad (8\text{-}11)$$

is a **large sample confidence interval** for μ, with confidence level of approximately $100(1-\alpha)\%$.

8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known

Example 8-4

EXAMPLE 8-4 Mercury Contamination

An article in the 1993 volume of the *Transactions of the American Fisheries Society* reports the results of a study to investigate the mercury contamination in largemouth bass. A sample of fish was selected from 53 Florida lakes, and mercury concentration in the muscle tissue, was measured (ppm). The mercury concentration values are

1.230	0.490	0.490	1.080	0.590	0.280	0.180	0.100	0.940
1.330	0.190	1.160	0.980	0.340	0.340	0.190	0.210	0.400
0.040	0.830	0.050	0.630	0.340	0.750	0.040	0.860	0.430
0.044	0.810	0.150	0.560	0.840	0.870	0.490	0.520	0.250
1.200	0.710	0.190	0.410	0.500	0.560	1.100	0.650	0.270
0.270	0.500	0.770	0.730	0.340	0.170	0.160	0.270	

8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known

Example 8-4 (continued)

The summary statistics from Minitab are displayed below:

Descriptive Statistics: Concentration

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Concentration	53	0.5250	0.4900	0.5094	0.3486	0.0479

Variable	Minimum	Maximum	Q1	Q3
Concentration	0.0400	1.3300	0.2300	0.7900

8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known

Example 8-4 (continued)

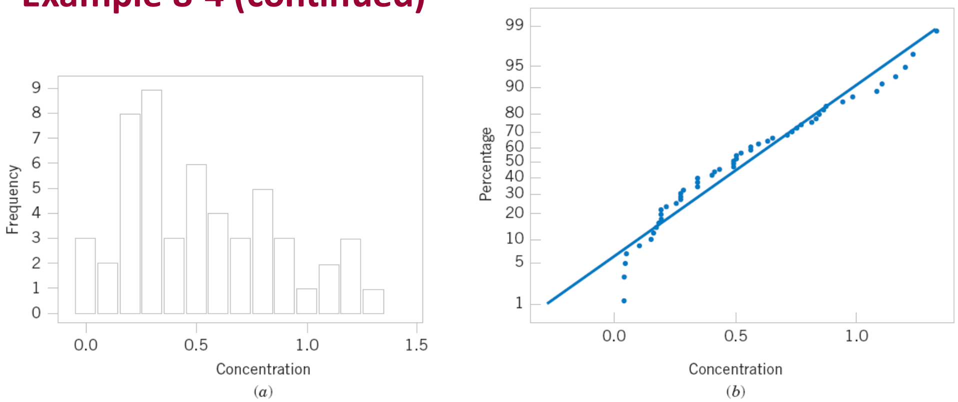

Figure 8-3 Mercury concentration in largemouth bass. (a) Histogram. (b) Normal probability plot.

Figure 8-3 Mercury concentration in largemouth bass (a) Histogram. (b) Normal probability plot

8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known

Example 8-4 (continued)

Figure 8-3(a) and (b) presents the histogram and normal probability plot of the mercury concentration data. Both plots indicate that the distribution of mercury concentration is not normal and is positively skewed. We want to find an approximate 95% CI on μ. Because $n > 40$, the assumption of normality is not necessary to use Equation 8-13. The required quantities are $n = 53$, $\bar{x} = 0.5250$, $s = 0.3486$, and $z_{0.025} = 1.96$. The approximate 95% CI on μ is

$$\bar{x} - z_{0.025}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + z_{0.025}\frac{s}{\sqrt{n}}$$

$$0.5250 - 1.96\frac{0.3486}{\sqrt{53}} \le \mu \le 0.5250 + 1.96\frac{0.3486}{\sqrt{53}}$$

$$0.4311 \le \mu \le 0.6189$$

This interval is fairly wide because there is a lot of variability in the mercury concentration measurements.

8-2 Confidence Interval on the Mean of a Normal Distribution, Variance Known

A General Large Sample Confidence Interval (Eq. 8-12)

$$\hat{\theta} - z_{\alpha/2}\sigma_{\hat{\Theta}} \le \theta \le \hat{\theta} + z_{\alpha/2}\sigma_{\hat{\Theta}} \qquad (8\text{-}12)$$

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

8-3.1 The *t* distribution (Eq. 8-13)

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with unknown mean μ and unknown variance σ^2. The random variable

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \qquad (8\text{-}13)$$

has a t distribution with $n - 1$ degrees of freedom.

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

8-3.1 The *t* distribution

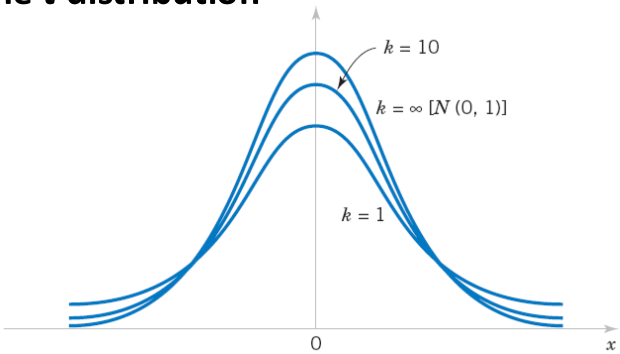

Figure 8-4 Probability density functions of several *t* distributions.

Figure 8-4 Probability density functions of several *t* distributions.

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

8-3.1 The *t* distribution

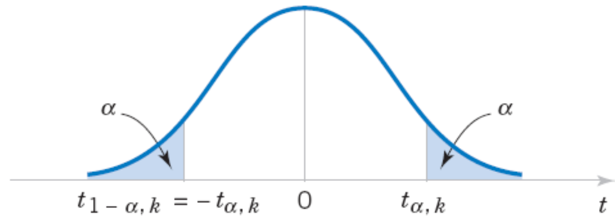

Figure 8-5 Percentage points of the *t* distribution.

Figure 8-5 Percentage points of the *t* distribution.

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

8-3.2 The *t* Confidence Interval on μ (Eq. 8-16)

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance σ^2, a **100(1 − α)% confidence interval on μ** is given by

$$\bar{x} - t_{\alpha/2,n-1}s/\sqrt{n} \le \mu \le \bar{x} + t_{\alpha/2,n-1}s/\sqrt{n} \qquad (8\text{-}16)$$

where $t_{\alpha/2,n-1}$ is the upper $100\alpha/2$ percentage point of the t distribution with $n-1$ degrees of freedom.

One-sided confidence bounds on the mean are found by replacing $t_{\alpha/2,n-1}$ in Equation 8-16 with $t_{\alpha,n-1}$.

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

Example 8-5

EXAMPLE 8-5 Alloy Adhesion

An article in the journal *Materials Engineering* (1989, Vol. II, No. 4, pp. 275–281) describes the results of tensile adhesion tests on 22 U-700 alloy specimens. The load at specimen failure is as follows (in megapascals):

19.8	10.1	14.9	7.5	15.4	15.4
15.4	18.5	7.9	12.7	11.9	11.4
11.4	14.1	17.6	16.7	15.8	
19.5	8.8	13.6	11.9	11.4	

The sample mean is $\bar{x} = 13.71$, and the sample standard deviation is $s = 3.55$. Figures 8-6 and 8-7 show a box plot and a normal probability plot of the tensile adhesion test data, respectively. These displays provide good support for the assumption that the population is normally distributed. We want to find a 95% CI on μ. Since $n = 22$, we have $n - 1 = 21$ degrees of freedom for t, so $t_{0.025,21} = 2.080$. The resulting CI is

$$\bar{x} - t_{\alpha/2,n-1}s/\sqrt{n} \le \mu \le \bar{x} + t_{\alpha/2,n-1}s/\sqrt{n}$$

$$13.71 - 2.080(3.55)/\sqrt{22} \le \mu \le 13.71 + 2.080(3.55)/\sqrt{22}$$

$$13.71 - 1.57 \le \mu \le 13.71 + 1.57$$

$$12.14 \le \mu \le 15.28$$

Practical Interpretation: The CI is fairly wide because there is a lot of variability in the tensile adhesion test measurements. A larger sample size would have led to a shorter interval.

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

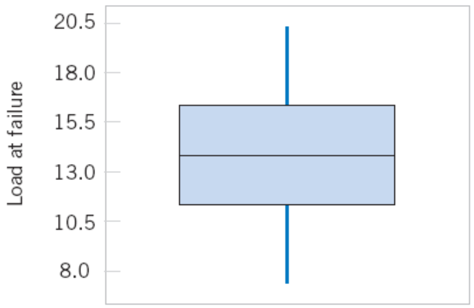

Figure 8-6 Box and whisker plot for the load at failure data in Example 8-5.

Figure 8-6 Box and Whisker plot for the load at failure data in Example 8-5.

8-3 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

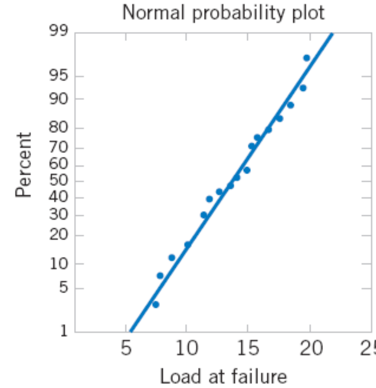

Figure 8-7 Normal probability plot of the load at failure data from Example 8-5.

Figure 8-7 Normal probability plot of the load at failure data in Example 8-5.

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

Definition (Eq. 8-17)

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2, and let S^2 be the sample variance. Then the random variable

$$X^2 = \frac{(n-1)\,S^2}{\sigma^2} \tag{8-17}$$

has a chi-square (χ^2) distribution with $n - 1$ degrees of freedom.

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

Figure 8-8 Probability density functions of several χ^2 distributions.

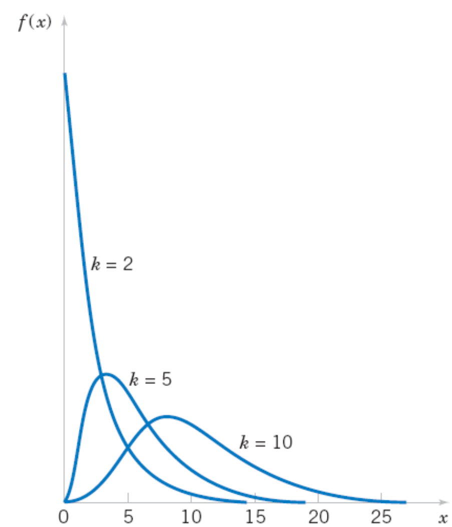

Figure 8-8 Probability density functions of several χ^2 distributions.

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

Definition

(Eq. 8-19)

If s^2 is the sample variance from a random sample of n observations from a normal distribution with unknown variance σ^2, then **a 100(1 − α)% confidence interval on σ^2 is**

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} \qquad (8\text{-}19)$$

where $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-\alpha/2,n-1}$ are the upper and lower $100\alpha/2$ percentage points of the chi-square distribution with $n-1$ degrees of freedom, respectively. A **confidence interval for σ** has lower and upper limits that are the square roots of the corresponding limits in Equation 8-19.

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

One-Sided Confidence Bounds

(Eq. 8-20)

The 100(1 − α)% lower and upper confidence bounds on σ^2 are

$$\frac{(n-1)s^2}{\chi^2_{\alpha,n-1}} \leq \sigma^2 \quad \text{and} \quad \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha,n-1}} \qquad (8\text{-}20)$$

respectively.

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

Example 8-6

EXAMPLE 8-6 Detergent Filling

An automatic filling machine is used to fill bottles with liquid detergent. A random sample of 20 bottles results in a sample variance of fill volume of $s^2 = 0.0153$ (fluid ounce)^2. If the variance of fill volume is too large, an unacceptable proportion of bottles will be under- or overfilled. We will assume that the fill volume is approximately normally distributed. A 95% upper confidence bound is found from Equation 8-26 as follows:

$$\sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{0.95,19}}$$

or

$$\sigma^2 \leq \frac{(19)0.0153}{10.117} = 0.0287 \text{ (fluid ounce)}^2$$

This last expression may be converted into a confidence interval on the standard deviation σ by taking the square root of both sides, resulting in

$$\sigma \leq 0.17$$

Practical Interpretation: Therefore, at the 95% level of confidence, the data indicate that the process standard deviation could be as large as 0.17 fluid ounce. The process engineer or manager now needs to determine if a standard deviation this large could lead to an operational problem with under-or over filled bottles.

8-5 A Large-Sample Confidence Interval For a Population Proportion

Normal Approximation for Binomial Proportion

If n is large, the distribution of

$$Z = \frac{X - np}{\sqrt{np(1-p)}} = \frac{\hat{P} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

is approximately standard normal.

The quantity $\sqrt{p(1-p)/n}$ is called the standard error of the point estimator $\hat{P}$.

8-5 A Large-Sample Confidence Interval For a Population Proportion

(Eq. 8-23)

If $\hat{p}$ is the proportion of observations in a random sample of size n that belongs to a class of interest, an approximate $100(1-\alpha)\%$ confidence interval on the proportion p of the population that belongs to this class is

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad (8\text{-}23)$$

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

8-5 A Large-Sample Confidence Interval For a Population Proportion

Example 8-7

EXAMPLE 8-7 Crankshaft Bearings

In a random sample of 85 automobile engine crankshaft bearings, 10 have a surface finish that is rougher than the specifications allow. Therefore, a point estimate of the proportion of bearings in the population that exceeds the roughness specification is $\hat{p} = x/n = 10/85 = 0.12$. A 95% two-sided confidence interval for p is computed from Equation 8-23 as

$$\hat{p} - z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

or

$$0.12 - 1.96\sqrt{\frac{0.12(0.88)}{85}} \le p \le 0.12 + 1.96\sqrt{\frac{0.12(0.88)}{85}}$$

which simplifies to

$$0.05 \le p \le 0.19$$

Practical Interpretation: This is a wide CI. While the sample size does not appear to be small ($n = 85$), the value of $\hat{p}$ is fairly small, which leads to a large standard error for $\hat{p}$ contributing to the wide CI.

8-5 A Large-Sample Confidence Interval For a Population Proportion

Choice of Sample Size

(Eq. 8-24 & 25)

The sample size for a specified value *E* is given by

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) \qquad (8\text{-}24)$$

An upper bound on *n* is given by

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 (0.25) \qquad (8\text{-}25)$$

8-5 A Large-Sample Confidence Interval For a Population Proportion

Example 8-8

EXAMPLE 8-8 Crankshaft Bearings

Consider the situation in Example 8-7. How large a sample is required if we want to be 95% confident that the error in using $\hat{p}$ to estimate p is less than 0.05? Using $\hat{p} = 0.12$ as an initial estimate of p, we find from Equation 8-24 that the required sample size is

$$n = \left(\frac{z_{0.025}}{E}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{1.96}{0.05}\right)^2 0.12(0.88) \cong 163$$

If we wanted to be *at least* 95% confident that our estimate $\hat{p}$ of the true proportion p was within 0.05 regardless of the value of p, we would use Equation 8-25 to find the sample size

$$n = \left(\frac{z_{0.025}}{E}\right)^2 (0.25) = \left(\frac{1.96}{0.05}\right)^2 (0.25) \cong 385$$

Practical Interpretation: Notice that if we have information concerning the value of p, either from a preliminary sample or from past experience, we could use a smaller sample while maintaining both the desired precision of estimation and the level of confidence.

8-5 A Large-Sample Confidence Interval For a Population Proportion

One-Sided Confidence Bounds

(Eq. 8-26)

The approximate $100(1-\alpha)\%$ lower and upper confidence bounds are

$$\hat{p} - z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \quad \text{and} \quad p \le \hat{p} + z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad (8\text{-}26)$$

respectively.

8-6 Guidelines for Constructing Confidence Intervals

Table 8-1 provides a simple road map to help select the appropriate analysis. Two primary comments can help identify the analysis:

1. Determine the parameter (and the distribution of the data) that will be bounded by the confidence interval or tested by the hypothesis.
2. Check if other parameters are known or need to be estimated.

8-7 Tolerance and Prediction Intervals

8-7.1 Prediction Interval for Future Observation

(Eq. 8-27)

A $100(1-\alpha)\%$ prediction interval (PI) on a single future observation from a normal distribution is given by

$$\bar{x} - t_{\alpha/2,n-1}\, s\sqrt{1+\frac{1}{n}} \le X_{n+1} \le \bar{x} + t_{\alpha/2,n-1}\, s\sqrt{1+\frac{1}{n}} \qquad (8\text{-}27)$$

The prediction interval for X_{n+1} will always be longer than the confidence interval for μ.

8-7 Tolerance and Prediction Intervals

Example 8-9

EXAMPLE 8-9 Alloy Adhesion

Reconsider the tensile adhesion tests on specimens of U-700 alloy described in Example 8-5. The load at failure for $n = 22$ specimens was observed, and we found that $\bar{x} = 13.71$ and $s = 3.55$. The 95% confidence interval on μ was $12.14 \le \mu \le 15.28$. We plan to test a twenty-third specimen. A 95% prediction interval on the load at failure for this specimen is

$$\bar{x} - t_{\alpha/2,n-1}\, s\sqrt{1+\frac{1}{n}} \le X_{n+1} \le \bar{x} + t_{\alpha/2,n-1}\, s\sqrt{1+\frac{1}{n}}$$

$$13.71 - (2.080)3.55\sqrt{1+\frac{1}{22}} \le X_{23} \le 13.71 + (2.080)3.55\sqrt{1+\frac{1}{22}}$$

$$6.16 \le X_{23} \le 21.26$$

Practical Interpretation: Notice that the prediction interval is considerably longer than the CI. This is because the CI an estimate of a parameter, while the PI is an interval estimate of a single future observation.

8-7 Tolerance and Prediction Intervals

8-7.2 Tolerance Interval for a Normal Distribution

Definition

A **tolerance interval** for capturing at least γ% of the values in a normal distribution with confidence level $100(1 - \alpha)$% is

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

where k is a tolerance interval factor found in Appendix Table XII. Values are given for γ = 90%, 95%, and 99% and for 90%, 95%, and 99% confidence.

8-7 Tolerance and Prediction Intervals

Example 8-10

EXAMPLE 8-10 Alloy Adhesion

Let's reconsider the tensile adhesion tests originally described in Example 8-5. The load at failure for $n = 22$ specimens was observed, and we found that $\bar{x} = 13.71$ and $s = 3.55$. We want to find a tolerance interval for the load at failure that includes 90% of the values in the population with 95% confidence. From Appendix Table XII, the tolerance factor k for $n = 22$, $\gamma = 0.90$, and 95% confidence is $k = 2.264$. The desired tolerance interval is

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

or

$$[13.71 - (2.264)3.55,\ 13.71 + (2.264)3.55]$$

which reduces to (5.67, 21.74).

Practical Interpretation: We can be 95% confident that at least 90% of the values of load at failure for this particular alloy lie between 5.67 and 21.74 megapascals.

Important Terms & Concepts of Chapter 8

Chi-squared distribution

Confidence coefficient

Confidence interval

Confidence interval for a:

- Population proportion
- Mean of a normal distribution
- Variance of a normal distribution

Confidence level

Error in estimation

Large sample confidence interval

1-sided confidence bounds

Precision of parameter estimation

Prediction interval

Tolerance interval

2-sided confidence interval

t distribution