CHAPTER 7

Confidence Intervals and Sample Size

Objectives

After completing this chapter, you should be able to

1. Find the confidence interval for the mean when σ is known.
2. Determine the minimum sample size for finding a confidence interval for the mean.
3. Find the confidence interval for the mean when σ is unknown.
4. Find the confidence interval for a proportion.
5. Determine the minimum sample size for finding a confidence interval for a proportion.
6. Find a confidence interval for a variance and a standard deviation.

Outline

Preview

This chapter presents the beginning of inferential statistics.

- The two major activities of inferential statistics are (1) to use sample data to estimate values of a population parameters, and(ch7). (2) to test hypotheses or claims made about population parameters(ch8).
- We introduce methods for estimating values of these important population parameters: means, proportions and variances.
- We also present methods for determining sample sizes necessary to estimate those parameters.

7.1 Confidence Intervals for the Mean When σ Is Known and Sample Size

- A point estimate is a specific numerical value estimate of a parameter.
- The best point estimate of the population mean μ is the sample mean $\bar{X}$. (unbiased estimator, not MD and mode)

Three Properties of a Good Estimator

1. The estimator should be an **unbiased estimator**. That is, the expected value or the mean of the estimates obtained from samples of a given size is equal to the parameter being estimated.

Three Properties of a Good Estimator

2. The estimator should be consistent. For a **consistent estimator**, as sample size increases, the value of the estimator approaches the value of the parameter estimated.

Three Properties of a Good Estimator

3. The estimator should be a **relatively efficient estimator**; that is, of all the statistics that can be used to estimate a parameter, the relatively efficient estimator has the smallest variance.

Confidence Intervals for the Mean When σ Is Known and Sample Size

- An **interval estimate** of a parameter is an interval or a range of values used to estimate the parameter.
- This estimate may or may not contain the value of the parameter being estimated.

Confidence Level of an Interval Estimate

- The **confidence level** of an interval estimate of a parameter is the probability that the interval estimate will contain the parameter, assuming that a large number of samples are selected and that the estimation process on the same parameter is repeated.

Confidence Interval

- A **confidence interval** is a specific interval estimate of a parameter determined by using data obtained from a sample and by using the specific confidence level of the estimate.

Formula for the Confidence Interval of the Mean for a Specific α

$$\bar{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$\bar{X}$: point estimate of μ

$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$: Maximum (marginal) error of estimate

$z_{\alpha/2}$: Critical value (C.V.), a *z* score can to distinguish between sample statistics that are likely to occur and those that are unlikely to occur from table E. α =1 – 95% C.L.

$\left(\frac{\sigma}{\sqrt{n}}\right)$: S.E.

For a 90% confidence interval: $z_{\alpha/2} = 1.65$

For a 95% confidence interval: $z_{\alpha/2} = 1.96$

For a 99% confidence interval: $z_{\alpha/2} = 2.58$

95% Confidence Interval of the Mean

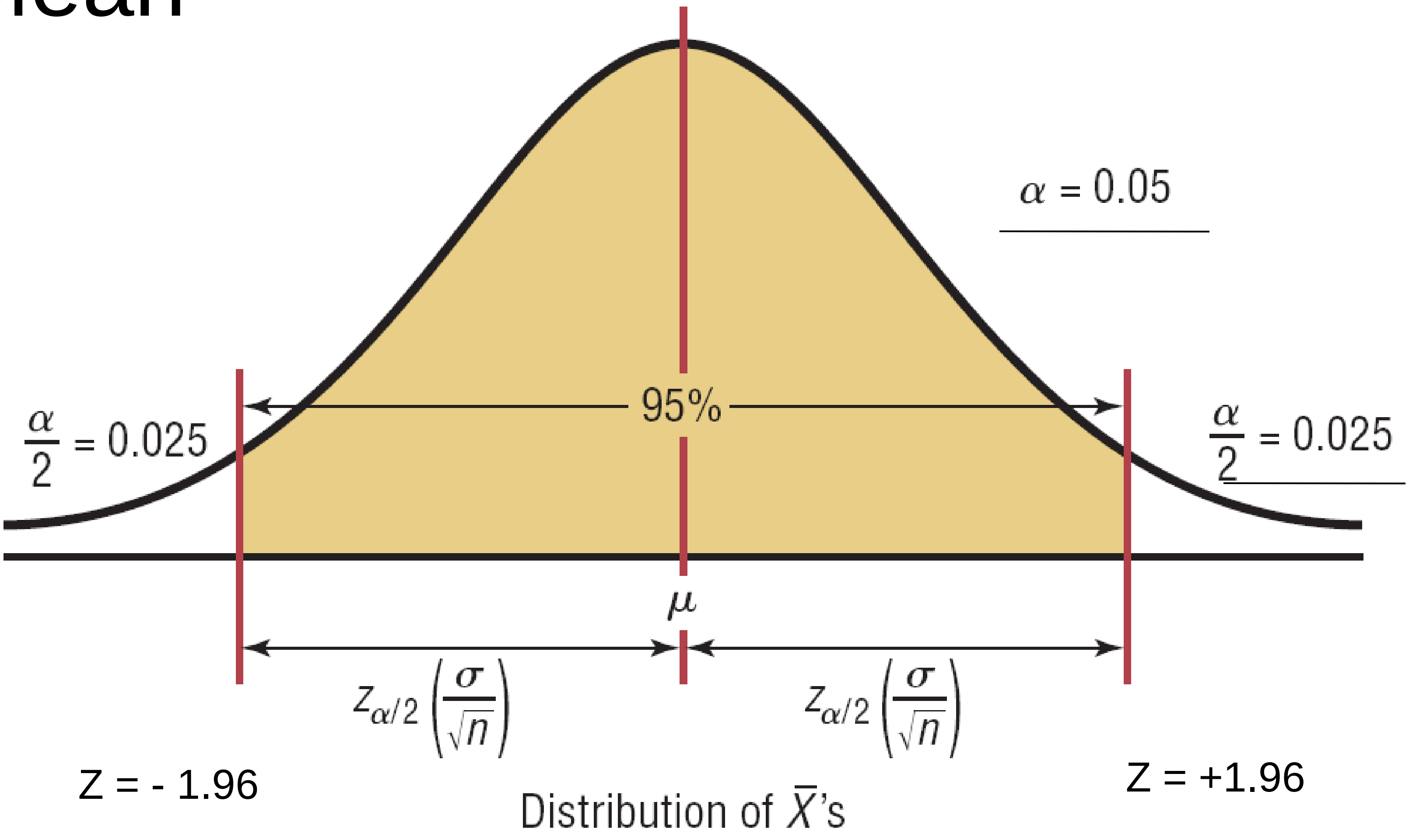


Table E The Standard Normal Distribution

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559

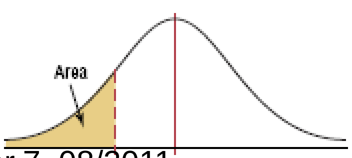
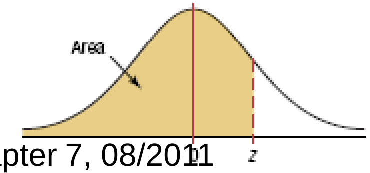


Table E *(continued)*

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857



Maximum Error of the Estimate

The maximum error of the estimate is the maximum likely difference between the point estimate of a parameter and the actual value of the parameter.

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Confidence Interval for a Mean

Rounding Rule

When you are computing a confidence interval for a population mean by using *raw data,* round off to one more decimal place than the number of decimal places in the original data.

When you are computing a confidence interval for a population mean by using a sample mean and a standard deviation, round off to the same number of decimal places as given for the mean.

Chapter 7
Confidence Intervals and Sample Size

Section 7-1

Example 7-1

Page #360

Example 7-1: Days to Sell an Aveo

A researcher wishes to estimate the number of days it takes an automobile dealer to sell a Chevrolet Aveo. A sample of 50 cars had a mean time on the dealer's lot of 54 days. Assume the **population standard deviation to be 6.0 days.** Find the best point estimate of the population mean and the 95% confidence interval of the population mean.

The best point estimate of the mean is 54 days.

$$\bar{X} = 54, \sigma = 6.0, n = 50, 95\% \rightarrow z = 1.96$$

$$\bar{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Example 7-1: Days to Sell an Aveo

$$\bar{X} = 54, \sigma = 6.0, n = 50, 95\% \rightarrow z = 1.96$$

$$\bar{X} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$54 - 1.96\left(\frac{6.0}{\sqrt{50}}\right) < \mu < 54 + 1.96\left(\frac{6.0}{\sqrt{50}}\right)$$

$$54 - 1.7 < \mu < 54 + 1.7$$

$$52.3 < \mu < 55.7$$

$$52 < \mu < 56$$

One can say with 95% confidence that the interval between 52 and 56 days contains the population mean, based on a sample of 50 automobiles.

Chapter 7
Confidence Intervals and Sample Size

Section 7-1

Example 7-2

Page #360

Example 7-2: Ages of Automobiles

A survey of 30 adults found that the mean age of a person's primary vehicle is 5.6 years. Assuming the **standard deviation of the population is 0.8 year**, find the best point estimate of the population mean and the 99% confidence interval of the population mean.

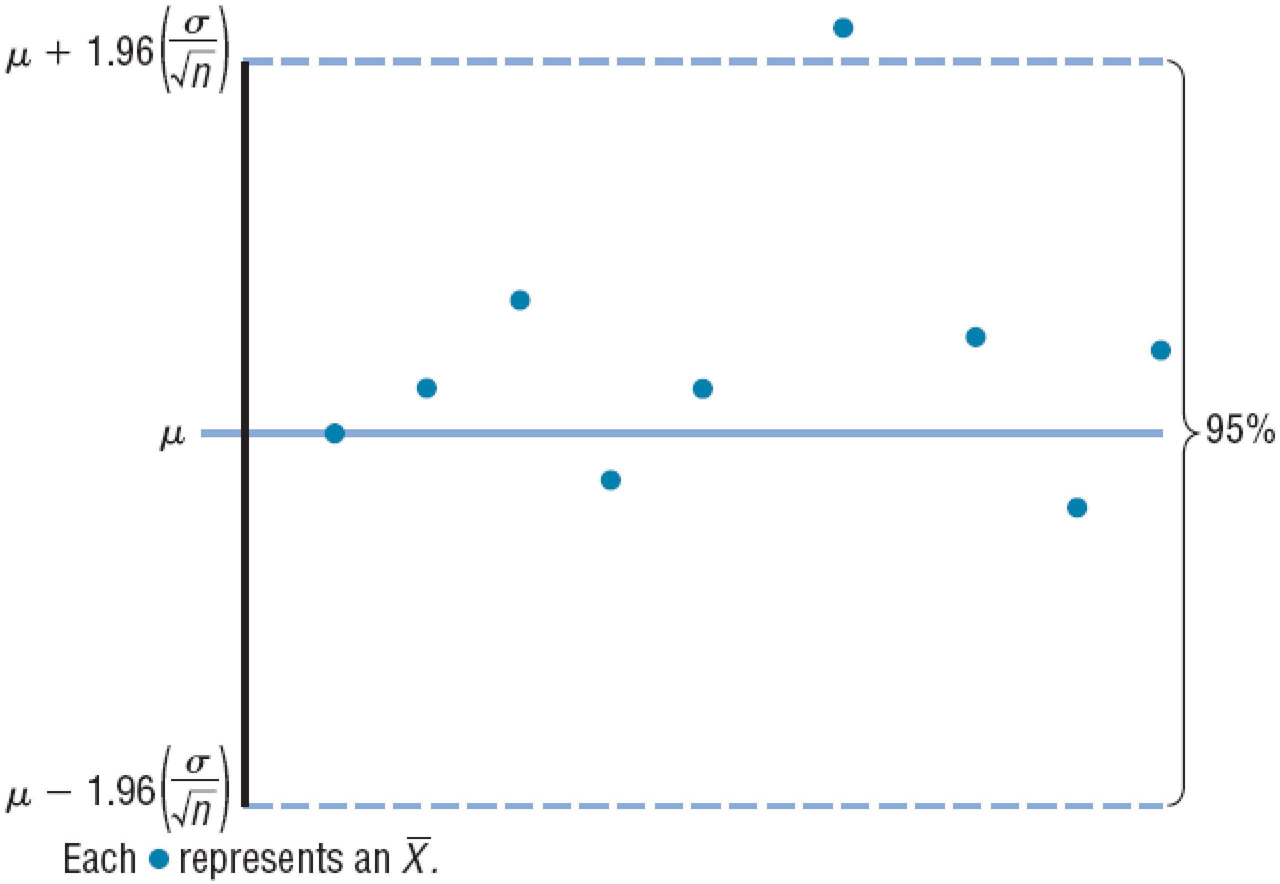
The best point estimate of the mean is 5.6 years.

$$5.6 - 2.58\left(\frac{0.8}{\sqrt{30}}\right) < \mu < 5.6 + 2.58\left(\frac{0.8}{\sqrt{30}}\right)$$

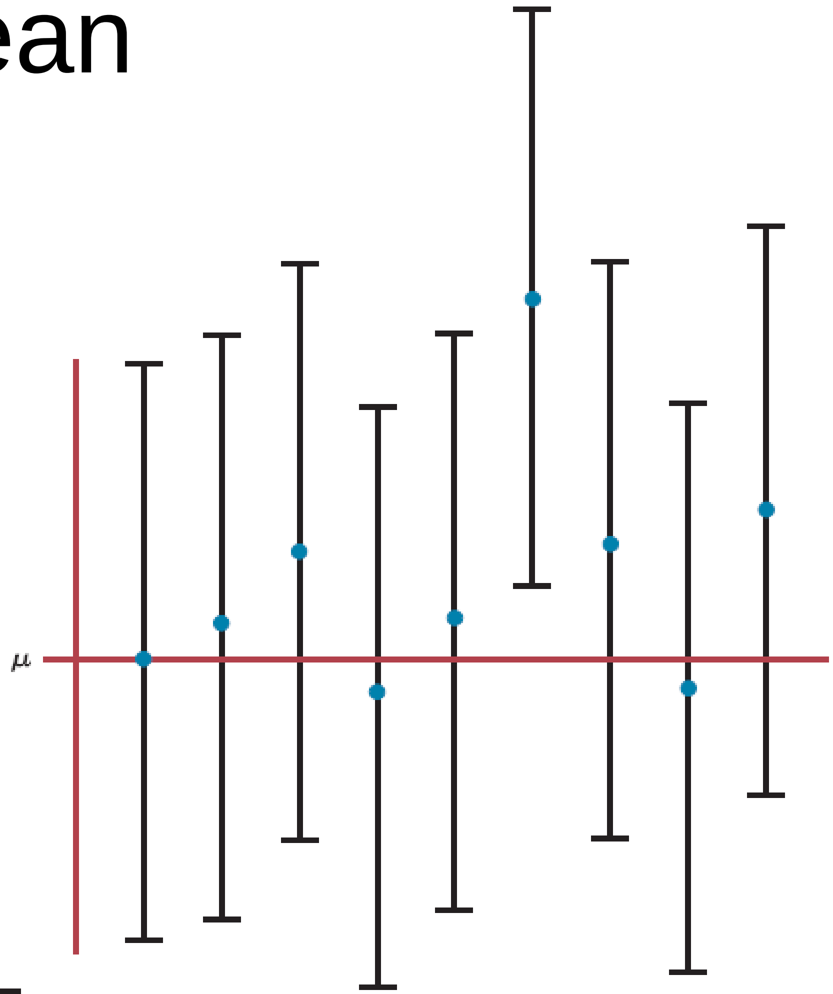
$$5.2 < \mu < 6.0$$

One can be 99% confident that the mean age of all primary vehicles is between 5.2 and 6.0 years, based on a sample of 30 vehicles.

95% Confidence Interval of the Mean



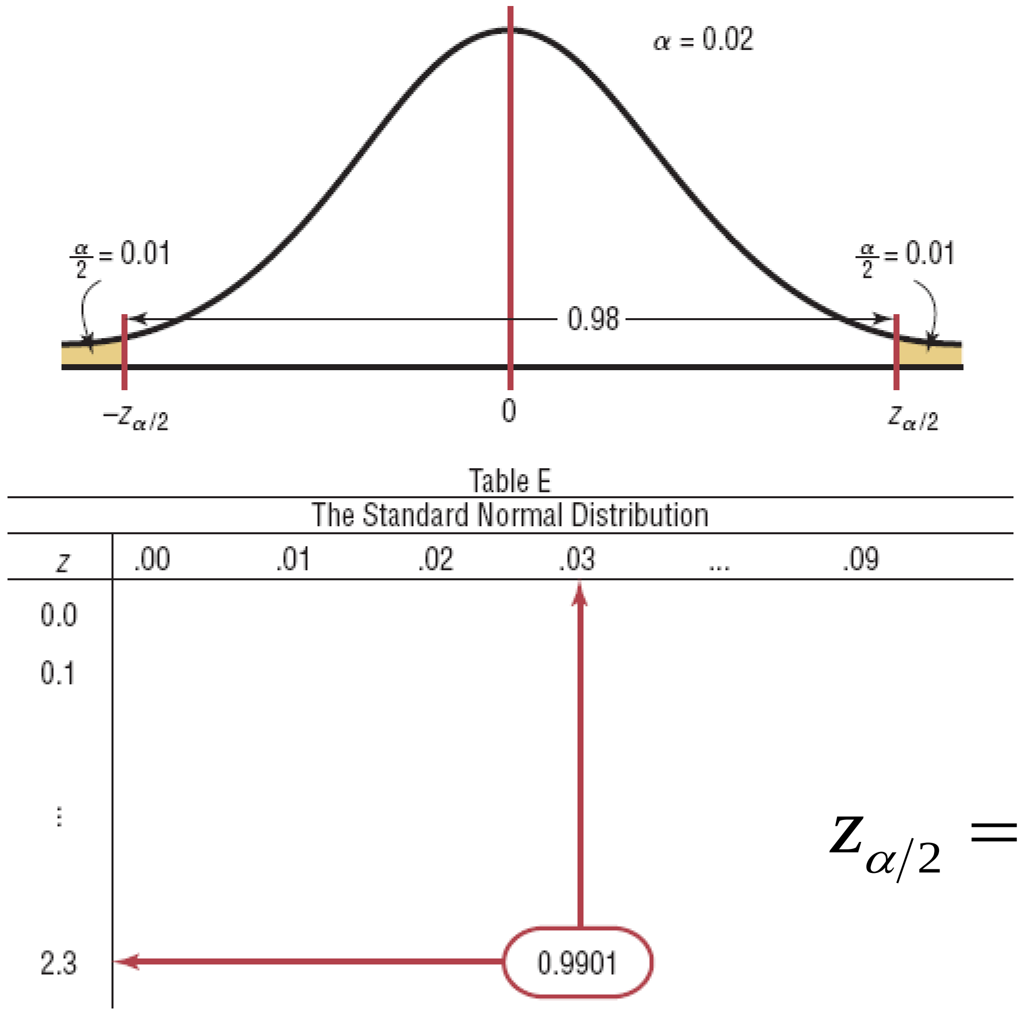
95% Confidence Interval of the Mean



Each I represents an interval about a sample mean.

One can be 95% confident that an interval built around a specific sample mean would contain the population mean.

Finding $z_{\alpha/2}$ for 98% CL. (option)



$$z_{\alpha/2} = 2.33$$

Technology Note (option)

This chapter and subsequent chapters include examples using raw data. If you are using computer or calculator programs to find the solutions, the answers you get may vary somewhat from the ones given in the textbook.

This is so because computers and calculators do not round the answers in the intermediate steps and can use 12 or more decimal places for computation. Also, they use more exact values than those given in the tables in the back of this book.

These discrepancies are part and parcel of statistics.

Formula for Minimum Sample Size Needed for an Interval Estimate of the Population Mean

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) \longrightarrow \qquad n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2$$

where *E* is the maximum error of estimate. If necessary, round the answer up to obtain a whole number. That is, if there is any fraction or decimal portion in the answer, use the next whole number for sample size *n.*

Chapter 7
Confidence Intervals and Sample Size

Section 7-1

Example 7-4

Page #364

Example 7-4: Depth of a River

A scientist wishes to estimate the average depth of a river. He wants to be 99% confident that the estimate is accurate within 2 feet. From a previous study, the standard deviation of the depths measured was 4.38 feet.

$$99\% \rightarrow z = 2.58, E = 2, \sigma = 4.38$$

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 = \left(\frac{2.58 \cdot 4.38}{2} \right)^2 = 31.92 = 32$$

Therefore, to be 99% confident that the estimate is within 2 feet of the true mean depth, the scientist needs at least a sample of 32 measurements.

7.2 Confidence Intervals for the Mean When σ Is Unknown

The value of σ, when it is not known, must be estimated by using *s*, the standard deviation of the sample.

When s is used, especially when the sample size is small (less than 30), critical values greater than the values for $z_{\alpha/2}$ are used in confidence intervals in order to keep the interval at a given level, such as the 95%.

These values are taken from the **Student t distribution**, most often called the **t distribution***.*

Characteristics of the *t* Distribution

The *t* distribution is similar to the standard normal distribution in these ways:

1. It is bell-shaped.
2. It is symmetric about the mean.
3. The mean, median, and mode are equal to 0 and are located at the center of the distribution.
4. The curve never touches the *x* axis.

Characteristics of the *t* Distribution

The *t* distribution differs from the standard normal distribution in the following ways:

1. The variance is greater than 1.
2. The *t* distribution is actually a family of curves based on the concept of degrees of freedom, which is related to sample size.
3. As the sample size increases, the *t* distribution approaches the standard normal distribution.

Degrees of Freedom

- The symbol d.f. will be used for **degrees of freedom**.
- The degrees of freedom for a confidence interval for the mean are found by subtracting 1 from the sample size. That is, d.f. = $n - 1$.
- Note: For some statistical tests used later in this book, the degrees of freedom are not equal to $n - 1$.

Formula for a Specific Confidence Interval for the Mean When σ Is Unknown

$$\bar{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{X} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

The degrees of freedom are *n - 1.*

Chapter 7
Confidence Intervals and Sample Size

Section 7-2

Example 7-5

Page #371

Example 7-5: Using Table F

Find the $t_{\alpha/2}$ value for a 95% confidence interval when the sample size is 22.

Degrees of freedom are d.f. = 21.

Table F

The t Distribution

	Confidence Intervals	50%	80%	90%	95%	98%	99%
d.f.	One tail α	0.25	0.10	0.05	0.025	0.01	0.005
	Two tails α	0.50	0.20	0.10	0.05	0.02	0.01
1							
2							
3							
⋮							
21					2.080	2.518	2.831
⋮							
$(z)\infty$		0.674	1.282[a]	1.645[b]	1.960	2.326[c]	2.576[d]

Chapter 7
Confidence Intervals and Sample Size

Section 7-2

Example 7-6

Page #372

Example 7-6: Sleeping Time

Ten randomly selected people were asked how long they slept at night. The mean time was 7.1 hours, and the standard deviation was 0.78 hour. Find the 95% confidence interval of the mean time. Assume the variable is normally distributed.

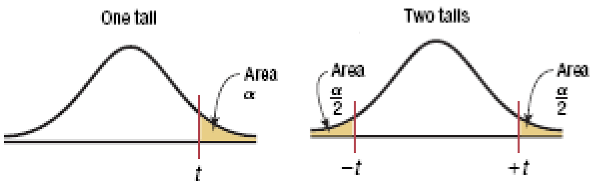
Since σ is unknown and *s* must replace it, the *t* distribution (Table F) must be used for the confidence interval. Hence, with 9 degrees of freedom, $t_{\alpha/2} = 2.262$.

$$\bar{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{X} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$7.1 - 2.262\left(\frac{0.78}{\sqrt{10}}\right) < \mu < 7.1 + 2.262\left(\frac{0.78}{\sqrt{10}}\right)$$

Table F The t Distribution

d.f.	Confidence intervals	80%	90%	95%	98%	99%
	One tail, α	0.10	0.05	0.025	0.01	0.005
	Two tails, α	0.20	0.10	0.05	0.02	0.01
1		3.078	6.314	12.706	31.821	63.657
2		1.886	2.920	4.303	6.965	9.925
3		1.638	2.353	3.182	4.541	5.841
4		1.533	2.132	2.776	3.747	4.604
5		1.476	2.015	2.571	3.365	4.032
6		1.440	1.943	2.447	3.143	3.707
7		1.415	1.895	2.365	2.998	3.499
8		1.397	1.860	2.306	2.896	3.355
9		1.383	1.833	2.262	2.821	3.250
10		1.372	1.812	2.228	2.764	3.169
11		1.363	1.796	2.201	2.718	3.106
12		1.356	1.782	2.179	2.681	3.055
13		1.350	1.771	2.160	2.650	3.012
14		1.345	1.761	2.145	2.624	2.977
15		1.341	1.753	2.131	2.602	2.947
16		1.337	1.746	2.120	2.583	2.921
17		1.333	1.740	2.110	2.567	2.898
18		1.330	1.734	2.101	2.552	2.878
19		1.328	1.729	2.093	2.539	2.861
20		1.325	1.725	2.086	2.528	2.845



Example 7-6: Sleeping Time

$$7.1-2.262\left(\frac{0.78}{\sqrt{10}}\right)<\mu<7.1+2.262\left(\frac{0.78}{\sqrt{10}}\right)$$

$$7.1-0.56<\mu<7.1+0.56$$

$$6.5<\mu<7.7$$

One can be 95% confident that the population mean is between 6.5 and 7.7 hours.

Chapter 7
Confidence Intervals and Sample Size

Section 7-2

Example 7-7

Page #372

Example 7-7: Home Fires by Candles

The data represent a sample of the number of home fires started by candles for the past several years. Find the 99% confidence interval for the mean number of home fires started by candles each year.

5460 5900 6090 6310 7160 8440 9930

Step 1: Find the mean and standard deviation.

The mean $\overline{X}$ = 7041.4 and

standard deviation s = 1610.3.

Step 2: Find $t_{\alpha/2}$ in Table F. The confidence level is 99%, and the degrees of freedom d.f. = 6

$t_{.005}$ = 3.707.

Example 7-7: Home Fires by Candles

Step 3: Substitute in the formula.

$$\bar{X} - t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{X} + t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

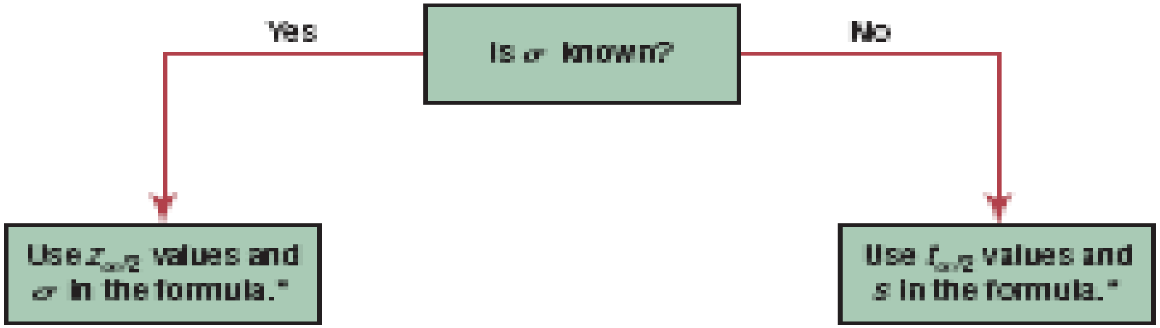
$$7041.4 - 3.707\left(\frac{1610.3}{\sqrt{7}}\right) < \mu < 7041.4 - 3.707\left(\frac{1610.3}{\sqrt{7}}\right)$$

$$7041.4 - 2256.2 < \mu < 7041.4 + 2256.2$$

$$4785.2 < \mu < 9297.6$$

One can be 99% confident that the population mean number of home fires started by candles each year is between 4785.2 and 9297.6, based on a sample of home fires occurring over a period of 7 years.

When to use the z or t distribution



*If $n < 30$, the variable must be normally distributed.

Finding the Point Estimate and *E* from a Confidence Interval

Point estimate of μ:

$$\bar{x} = \frac{\text{(upper confidence limit) + (lower confidence limit)}}{2}$$

Margin of Error:

$$E = \frac{\text{(upper confidence limit) – (lower confidence limit)}}{2}$$

7.3 Confidence Intervals and Sample Size for Proportions

p = population proportion

$\hat{p}$ (read *p* "hat") = sample proportion

For a sample proportion,

$$\hat{p} = \frac{X}{n} \quad \text{and} \quad \hat{q} = \frac{n - X}{n} \quad or \quad \hat{q} = 1 - \hat{p}$$

where X = number of sample units that possess the characteristics of interest and n = sample size.

Chapter 7
Confidence Intervals and Sample Size

<u>Section 7-3</u>

Example 7-8

Page #378

Example 7-8: Air Conditioned Households

In a recent survey of 150 households, 54 had central air conditioning. Find $\hat{p}$ and $\hat{q}$, where $\hat{p}$ is the proportion of households that have central air conditioning.

Since X = *54 and* n = *150,*

$$\hat{p} = \frac{X}{n} = \frac{54}{150} = 0.36 = 36\%$$

$$\hat{q} = 1 - \hat{p} = 1 - 0.36 = 0.64 = 64\%$$

Formula for a Specific Confidence Interval for a Proportion

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

when $np \geq 5$ and $nq \geq 5$.

Rounding Rule: Round off to three decimal places.

Chapter 7
Confidence Intervals and Sample Size

Section 7-3

Example 7-9

Page #378

Example 7-9: Male Nurses

A sample of 500 nursing applications included 60 from men. Find the 90% confidence interval of the true proportion of men who applied to the nursing program.

$$\hat{p} = X/n = 60/500 = 0.12, \ \ \hat{q} = 0.88$$

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.12 - 1.65\sqrt{\frac{(0.12)(0.88)}{500}} < p < 0.12 + 1.65\sqrt{\frac{(0.12)(0.88)}{500}}$$

$$0.12 - 0.024 < p < 0.12 + 0.024$$

$$.096 < p < 0.144$$

You can be 90% confident that the percentage of applicants who are men is between 9.6% and 14.4%.

Chapter 7
Confidence Intervals and Sample Size

Section 7-3

Example 7-10

Page #379

Example 7-10: Religious Books

A survey of 1721 people found that 15.9% of individuals purchase religious books at a Christian bookstore. Find the 95% confidence interval of the true proportion of people who purchase their religious books at a Christian bookstore.

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.159 - 1.96\sqrt{\frac{(0.159)(0.841)}{1721}} < p < 0.159 + 1.96\sqrt{\frac{(0.159)(0.841)}{1721}}$$

$$0.142 < p < 0.176$$

You can say with 95% confidence that the true percentage is between 14.2% and 17.6%.

Formula for Minimum Sample Size Needed for Interval Estimate of a Population Proportion

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

If necessary, round up to the next whole number.

Chapter 7
Confidence Intervals and Sample Size

Section 7-3

Example 7-11

Page #380

Example 7-11: Home Computers

A researcher wishes to estimate, with 95% confidence, the proportion of people who own a home computer. A previous study shows that 40% of those interviewed had a computer at home. The researcher wishes to be accurate within 2% of the true proportion. Find the minimum sample size necessary.

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 = (0.40)(0.60)\left(\frac{1.96}{0.02}\right)^2 = 2304.96$$

The researcher should interview a sample of at least 2305 people.

Chapter 7
Confidence Intervals and Sample Size

Section 7-3

Example 7-12

Page #380

Example 7-12: Car Phone Ownership

The same researcher wishes to estimate the proportion of executives who own a car phone. She wants to be 90% confident and be accurate within 5% of the true proportion. Find the minimum sample size necessary.

Since there is no prior knowledge of $\hat{p}$, statisticians assign the values $\hat{p}$ = 0.5 and $\hat{q}$ = 0.5. The sample size obtained by using these values will be large enough to ensure the specified degree of confidence.

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 = (0.50)(0.50)\left(\frac{1.65}{0.05}\right)^2 = 272.25$$

The researcher should ask at least 273 executives.

7-4 Confidence Intervals for Variances and Standard Deviations

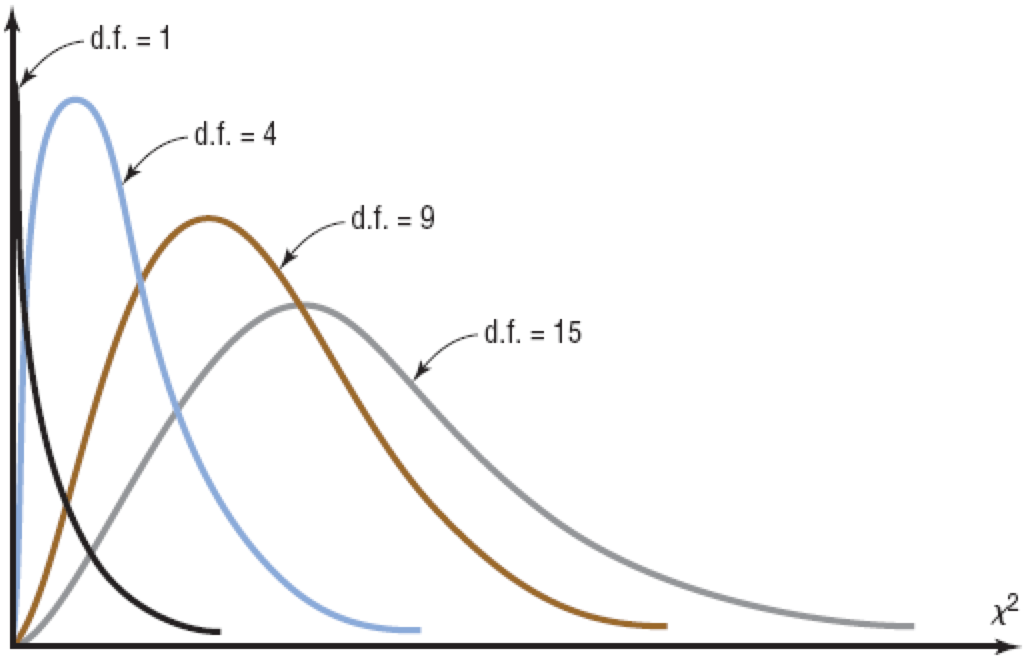
- When products that fit together (such as pipes) are manufactured, it is important to keep the variations of the diameters of the products as small as possible; otherwise, they will not fit together properly and will have to be scrapped.
- In the manufacture of medicines, the variance and standard deviation of the medication in the pills play an important role in making sure patients receive the proper dosage.
- For these reasons, confidence intervals for variances and standard deviations are necessary.

Chi-Square Distributions

- The chi-square distribution must be used to calculate confidence intervals for variances and standard deviations.
- The chi-square variable is similar to the *t* variable in that its distribution is a family of curves based on the number of degrees of freedom.
- The symbol for chi-square is χ^2 (Greek letter chi, pronounced "ki").
- A chi-square variable cannot be negative, and the distributions are skewed to the right.

Chi-Square Distributions

- At about 100 degrees of freedom, the chi-square distribution becomes somewhat symmetric.
- The area under each chi-square distribution is equal to 1.00, or 100%.



Formula for the Confidence Interval for a Variance

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}, \qquad \text{d.f.} = n-1$$

Formula for the Confidence Interval for a Standard Deviation

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{right}}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{\text{left}}}}, \qquad \text{d.f.} = n-1$$

Chapter 7
Confidence Intervals and Sample Size

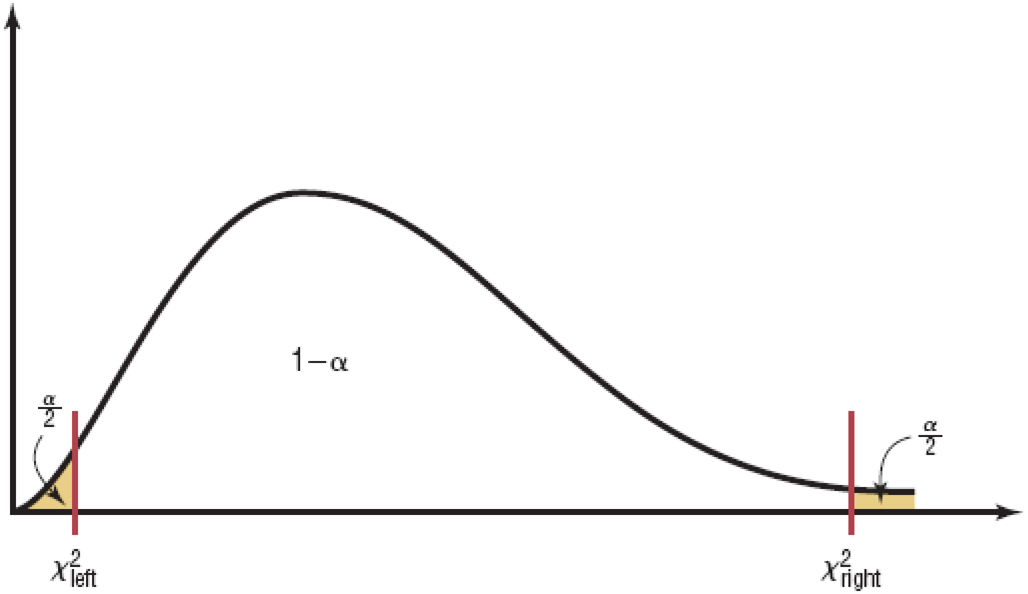
Section 7-4

Example 7-13

Page #387

Example 7-13: Using Table G

Find the values for χ^2_{right} and χ^2_{left} for a 90% confidence interval when *n = 25.*



To find χ^2_{right}, subtract 1 - 0.90 = 0.10. Divide by 2 to get 0.05.

To find χ^2_{left} , subtract 1 - 0.05 to get 0.95.

Example 7-13: Using Table G

Use the 0.95 and 0.05 columns and the row corresponding to 24 d.f. in Table G.

Table G

The Chi-square Distribution

Degrees of freedom	α 0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1										
2										
⋮										
24				13.848			36.415			
				χ^2_{left}			χ^2_{right}			

The χ^2_{right} value is 36.415; the χ^2_{left} value is 13.848.

Area represents the cumulative area located to the right of the critical value.

Confidence Interval for a Variance or Standard Deviation

Rounding Rule

When you are computing a confidence interval for a population variance or standard deviation by using raw data, round off to one more decimal places than the number of decimal places in the original data.

When you are computing a confidence interval for a population variance or standard deviation by using a sample variance or standard deviation, round off to the same number of decimal places as given for the sample variance or standard deviation.

Chapter 7
Confidence Intervals and Sample Size

Section 7-4

Example 7-14

Page #389

Example 7-14: Nicotine Content

Find the 95% confidence interval for the variance and standard deviation of the nicotine content of cigarettes manufactured if a sample of 20 cigarettes has a standard deviation of 1.6 milligrams.

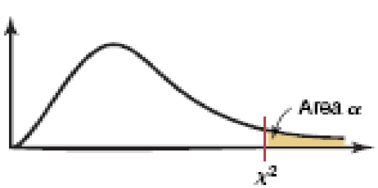
To find χ^2_{right}, subtract 1 - 0.95 = 0.05. Divide by 2 to get 0.025.

To find χ^2_{left}, subtract 1 - 0.025 to get 0.975.

In Table G, the 0.025 and 0.975 columns with the d.f. 19 row yield values of 32.852 and 8.907, respectively.

Table G The Chi-Square Distribution

Degrees of freedom	α									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997



Example 7-14: Nicotine Content

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

$$\frac{(19)(1.6)^2}{32.852} < \sigma^2 < \frac{(19)(1.6)^2}{8.907}$$

$$1.5 < \sigma^2 < 5.5$$

You can be 95% confident that the true variance for the nicotine content is between 1.5 and 5.5 milligrams.

$$\sqrt{1.5} < \sigma < \sqrt{5.5}$$

$$1.2 < \sigma < 2.3$$

You can be 95% confident that the true standard deviation is between 1.2 and 2.3 milligrams.

Chapter 7
Confidence Intervals and Sample Size

Section 7-4

Example 7-15

Page #389

Example 7-15: Cost of Ski Lift Tickets

Find the 90% confidence interval for the variance and standard deviation for the price in dollars of an adult single-day ski lift ticket. The data represent a selected sample of nationwide ski resorts. Assume the variable is normally distributed.

59	54	53	52	51
39	49	46	49	48

Using technology, we find the variance of the data is $s^2 = 28.2$.

In Table G, the 0.05 and 0.95 columns with the d.f. 9 row yield values of 16.919 and 3.325, respectively.

Example 7-15: Cost of Ski Lift Tickets

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

$$\frac{(9)(28.2)}{16.919} < \sigma^2 < \frac{(9)(28.2)}{3.325}$$

$$15.0 < \sigma^2 < 76.3$$

You can be 95% confident that the true variance for the cost of ski lift tickets is between 15.0 and 76.3.

$$\sqrt{15.0} < \sigma < \sqrt{76.3}$$

$$3.87 < \sigma < 8.73$$

You can be 95% confident that the true standard deviation is between \$3.87 and \$8.73.