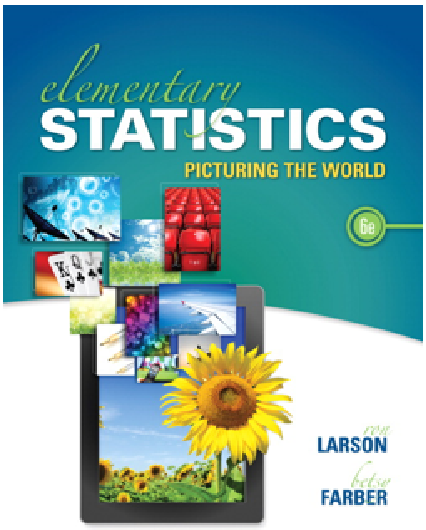
Elementary Statistics: Picturing The World

Sixth Edition



Chapter 6

Confidence Intervals

6.1 pg. 305 5-21 odd,24,25,26,29,32,37,39,47,50

6.2 pg. 315 5,6,9,13,17,21,26 or 27,30

6.3 pg. 325 4,5,7,8,11,13,17,18,23

6.4 pg. 334 3,4,7,9,10,15,19,21

Statistics Chapter 6

6.1 Condidence intervals for the MEAN (when σ is known)

DEFINITION

A **point estimate** is a single value estimate for a population parameter. The most unbiased point estimate of the population mean μ is the sample mean $\bar{x}$.

The validity of an estimation method is increased when you use a sample statistic that is unbiased and has low variability. A statistic is unbiased if it does not overestimate or underestimate the population parameter. In Chapter 5, you learned that the mean of all possible sample means of the same size equals the population mean. As a result, $\bar{x}$ is an **unbiased estimator** of μ. When the standard error $\sigma / \sqrt{n}$ of a sample mean is decreased by increasing n, it becomes less variable.

EXAMPLE 1

Finding a Point Estimate

An economics researcher is collecting data about grocery store employees in a county. The data listed below represents a random sample of the number of hours worked by 40 employees from several grocery stores in the county. Find a point estimate of the population mean μ.

30	26	33	26	26	33	31	31	21	37
27	20	34	35	30	24	38	34	39	31
22	30	23	23	31	44	31	33	33	26
27	28	25	35	23	32	29	31	25	27

Number of hours					
26	25	32	31	28	28
28	22	28	25	21	40
32	22	25	22	26	24
46	20	35	22	32	48
32	36	38	32	22	19

Try It Yourself 1

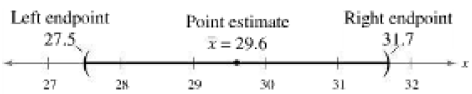
Another random sample of the hours worked by 30 grocery store employees in the county is shown at the left. Use this sample to find another point estimate for μ.

a. Find the sample mean.
b. Estimate the population mean.

DEFINITION

An **interval estimate** is an interval, or range of values, used to estimate a population parameter.

Although you can assume that the point estimate in Example 1 is not equal to the actual population mean, it is probably close to it. To form an interval estimate, use the point estimate as the center of the interval, and then add and subtract a margin of error. For instance, if the margin of error is 2.1, then an interval estimate would be given by 29.6 ± 2.1 or $27.5 < \mu < 31.7$. The point estimate and interval estimate are shown in the figure.



Interval Estimate

Before finding a margin of error for an interval estimate, you should first determine how confident you need to be that your interval estimate contains the population mean μ.

DEFINITION

The **level of confidence** c is the probability that the interval estimate contains the population parameter, assuming that the estimation process is repeated a large number of times.

Chapter 6: Confidence Intervals

$\overline{x}$ μ

Lesson 6-1: CIs for the Mean - LARGE SAMPLES

This section is all about when n is AT LEAST 30.*

We have been using $\overline{x}$ as an approximation for the true population mean μ. How accurate is this? How CONFIDENT are we that we are close to the true mean?

This is what a confidence interval tells us: "We are 95% (90%, 99%, whatever) confident that the true mean is between a and b

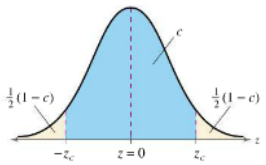
Study Tip

In this course, you will usually use 90%, 95%, and 99% levels of confidence. Here are the z-scores that correspond to these levels of confidence.

Level of Confidence	z_c
90%	1.645
95%	1.96
99%	2.575

Pay Attention to this! Save yourself lots of time looking it up constantly!!!!!!

Otherwise, you have to do this every time:



If $c = 90\%$:	
$c = 0.90$	Area in blue region
$1 - c = 0.10$	Area in yellow regions
$\frac{1}{2}(1 - c) = 0.05$	Area in one tail
$-z_c = -1.645$	Critical value separating left tail
$z_c = 1.645$	Critical value separating right tail

The difference between a point estimate (like $\overline{x}$) and the true parameter value (for instance μ) is the SAMPLING ERROR. The margin of error is how much "wiggle room" we have to allow for.

DEFINITION

Given a level of confidence c, the **margin of error** E (sometimes also called the maximum error of estimate or error tolerance) is the greatest possible distance between the point estimate and the value of the parameter it is estimating. For a population mean μ where σ is known, the margin of error is

$$E = z_c \sigma_{\overline{x}} = z_c \frac{\sigma}{\sqrt{n}} \quad \text{Margin of error for } \mu \ (\sigma \text{ known})$$

when these conditions are met.

1. The sample is random.
2. At least one of the following is true: The population is normally distributed or $n \geq 30$.

EXAMPLE 2: Finding the Margin of Error

Finding the Margin of Error

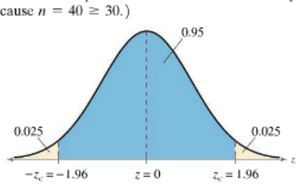
Use the data in Example 1 and a 95% confidence level to find the margin of error for the mean number of hours worked by grocery store employees. Assume the population standard deviation is 7.9 hours.

Solution:

Because σ is known ($\sigma = 7.9$), the sample is random (see Example 1), and $n = 40 \geq 30$, use the formula for E given above. The z-score that corresponds to a 95% confidence level is 1.96. This implies that 95% of the area under the standard normal curve falls within 1.96 standard deviations of the mean. (You can approximate the distribution of the sample means with a normal curve by the Central Limit Theorem because $n = 40 \geq 30$.)

Using the values $z_c = 1.96$, $\sigma = 7.9$, and $n = 40$,

$$E = z_c \frac{\sigma}{\sqrt{n}}$$
$$= 1.96 \cdot \frac{7.9}{\sqrt{40}}$$
$$\approx 2.4.$$



Interpretation You are 95% confident that the margin of error for the population mean is about 2.4 hours.

Lesson 6-1 (continued)

$\bar{x}$ μ

Using the point estimate $\bar{x}$ and the margin of error E, we can make a CONFIDENCE INTERVAL for the true mean μ as follows:

DEFINITION

A ***c*-confidence interval** for the population mean μ is

$$\bar{x} - E < \mu < \bar{x} + E.$$

The probability that the confidence interval contains μ is c.

Our final answer can be written as: (fill in the blanks)

"We are ____% confident that the true mean is between _____ and _______."

The general process - ONLY IF N > 30!!!

EXAMPLE 3

See Minitab steps on page 344.

Constructing a Confidence Interval

Use the data in Example 1 to construct a 95% confidence interval for the mean number of hours worked by grocery store employees.

30	26	33	26	26	33	31	31	21	37
27	20	34	35	30	24	38	34	39	31
22	30	23	23	31	44	31	33	33	26
27	28	25	35	23	32	29	31	25	27

GUIDELINES

Constructing a Confidence Interval for a Population Mean (σ Known)

IN WORDS	IN SYMBOLS
1. Verify that σ is known, the sample is random, and either the population is normally distributed or $n \geq 30$.	
2. Find the sample statistics n and $\bar{x}$.	$\bar{x} = \frac{\Sigma x}{n}$
3. Find the critical value z_c that corresponds to the given level of confidence.	Use Table 4 in Appendix B.
4. Find the margin of error E.	$E = z_c \frac{\sigma}{\sqrt{n}}$
5. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\bar{x} - E$ Right endpoint: $\bar{x} + E$ Interval: $\bar{x} - E < \mu < \bar{x} + E$

6. Write the interpretation sentence

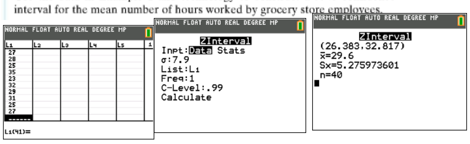
"We are ____% confident that the average number of hours worked is between _____ and _______."

EXAMPLE 4: Constructing a Confidence Interval

EXAMPLE 4

Constructing a Confidence Interval Using Technology

Use the data in Example 1 and technology to construct a 99% confidence interval for the mean number of hours worked by grocery store employees.



Study Tip

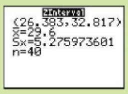
Using a TI-84 Plus, you can either enter the original data into a list to construct the confidence interval or enter the descriptive statistics.

STAT

Choose the TESTS menu.

7: ZInterval...

Select the *Data* input option when you use the original data. Select the *Stats* input option when you use the descriptive statistics. In each case, enter the appropriate values, then select *Calculate*. Your results may differ slightly depending on the method you use. For Example 4, the original data were entered.



Example 5

A college admissions director wishes to estimate the mean age of all students currently enrolled. In a random sample of 20 students, the mean age is found to be 22.9 years. From past studies, the standard deviation is known to be 1.5 years, and the population is normally distributed. Construct a 90% confidence interval of the students' mean age.

$\bar{x}$

Solution:

"We are ____% confident that the average age of students is between _____ and _______."

Lesson 6-1 (continued) - Sample Size

As the level of confidence decreases, the interval gets wider.

As the interval widens, the precision of the estimate decreases.

We can fix this by increasing the sample size.

To find the sample size needed to reach the desired level of confidence:

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

FIND A MINIMUM SAMPLE SIZE TO ESTIMATE μ

Given a c-confidence level and a margin of error E, the minimum sample size n needed to estimate the population mean μ is

$$n = \left(\frac{z_c \sigma}{E}\right)^2.$$

If σ is unknown, you can estimate it using s, provided you have a preliminary sample with at least 30 members.

EXAMPLE 6: Determining the minimum sample size

EXAMPLE 6

Determining a Minimum Sample Size

The economics researcher in Example 1 wants to estimate the mean number of hours worked by all grocery store employees in the county. How many employees must be included in the sample to be 95% confident that the sample mean is within 1.5 hours of the population mean?

pg 317: 5-21 odd, 24, 25, 26,29 32, 37, 39, 47, 50

Chapter 6: Confidence Intervals

Lesson 6-2: CIs for the Mean - SMALL SAMPLES

When n is less than 30, we have to use the t-distribution

We used the normal distribution last section, since we had at least 30 samples.

Sometimes it isn't practical to get 30 or more samples

- rare diseases' treatments

- limited pool to draw from (maybe there are only 15 schools willing to admit they teach underwater basket weaving)

The *t*-distribution:

DEFINITION

If the distribution of a random variable *x* is approximately normal, then

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

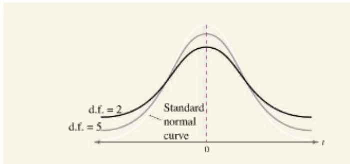
follows a ***t*-distribution.**

Critical values of *t* are denoted by t_c. Here are several properties of the *t*-distribution.

1. The mean, median, and mode of the *t*-distribution are equal to 0.
2. The *t*-distribution is bell-shaped and symmetric about the mean.
3. The total area under the *t*-distribution curve is equal to 1.
4. The tails in the *t*-distribution are "thicker" than those in the standard normal distribution.
5. The standard deviation of the *t*-distribution varies with the sample size, but it is greater than 1.
6. The *t*-distribution is a family of curves, each determined by a parameter called the *degrees of freedom*. The **degrees of freedom** (sometimes abbreviated as d.f.) are the number of free choices left after a sample statistic such as $\bar{x}$ is calculated. When you use a *t*-distribution to estimate a population mean, the degrees of freedom are equal to one less than the sample size.

   d.f. $= n - 1$ Degrees of freedom

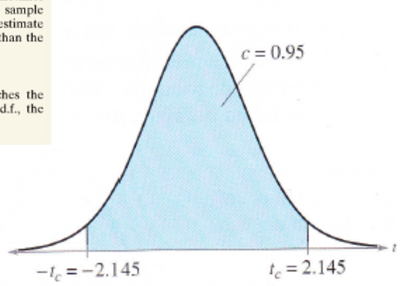
7. As the degrees of freedom increase, the *t*-distribution approaches the standard normal distribution, as shown in the figure. After 30 d.f., the *t*-distribution is close to the standard normal distribution.



Using a t-table:

Notation: $t_{df,\,c}$

For example:



The circled value corresponds to $t_{14,\,0.95}$

	Level of confidence, c	0.50	0.80	0.90	0.95	0.98
	One tail, α	0.25	0.10	0.05	0.025	0.01
d.f.	Two tails, α	0.50	0.20	0.10	0.05	0.02
1		1.000	3.078	6.314	12.706	31.821
2		.816	1.886	2.920	4.303	6.965
3		.765	1.638	2.353	3.182	4.541
12		.695	1.356	1.782	2.179	2.681
13		.694	1.350	1.771	2.160	2.650
14		.692	1.345	1.761	2.145	2.624
15		.691	1.341	1.753	2.131	2.602
16		.690	1.337	1.746	2.120	2.583
28		.683	1.313	1.701	2.048	2.467
29		.683	1.311	1.699	2.045	2.462
∞		.674	1.282	1.645	1.960	2.326

Example 1: Finding critical values of t

t_c

Find the t-value for a 95% confidence level when the sample size is 15.

Solution:

Lesson 6-2 (continued)

Constructing a confidence interval for samples under 30

GUIDELINES

Constructing a Confidence Interval for a Population Mean (σ Unknown)

IN WORDS	IN SYMBOLS
1. Verify that σ is not known, the sample is random, and either the population is normally distributed or $n \geq 30$.	
2. Find the sample statistics n, $\bar{x}$, and s.	$\bar{x} = \frac{\Sigma x}{n}$, $s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$
3. Identify the degrees of freedom, the level of confidence c, and the critical value t_c.	d.f. $= n - 1$ Use Table 5 in Appendix B.
4. Find the margin of error E.	$E = t_c \frac{s}{\sqrt{n}}$
5. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\bar{x} - E$ Right endpoint: $\bar{x} + E$ Interval: $\bar{x} - E < \mu < \bar{x} + E$

Study Tip

Remember that you can calculate the sample standard deviation s using the formula

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

or the shortcut formula

$$s = \sqrt{\frac{\Sigma x^2 - (\Sigma x)^2/n}{n-1}}.$$

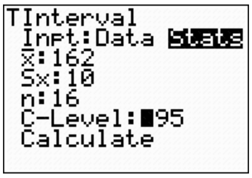
However, the most convenient way to find the sample standard deviation is to use the *1-Var* Stats feature of a graphing calculator.

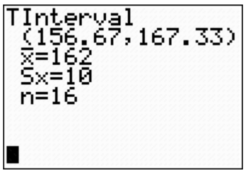
Example 2: Constructing a Confidence Interval (Small sample)

You randomly select 16 coffee shops and measure the temperature of the coffee sold at each. The sample mean temperature is 162.0 F with a sample standard deviation os 10.0 F. Find the 95% confidence interval for the mean temperature. Assume the temperatures are normally distributed.

Solution:

"We are ____% confident that the average coffee temperature is between _____ and ______."



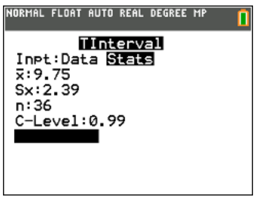


EXAMPLE 3

See TI-84 Plus steps on page 345.

Constructing a Confidence Interval

You randomly select 36 cars of the same model that were sold at a car dealership and determine the number of days each car sat on the dealership's lot before it was sold. The sample mean is 9.75 days, with a sample standard deviation of 2.39 days. Construct a 99% confidence interval for the population mean number of days the car model sits on the dealership's lot.



"We are ____% confident that the population mean number of days the car model sits on the dealership's lot is between _____ and ______."

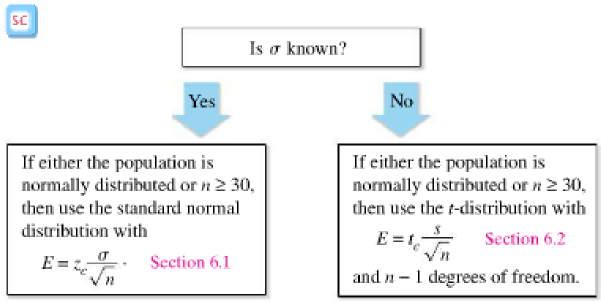
Try It Yourself 3

Construct 90% and 95% confidence intervals for the population mean number of days the car model sits on the dealership's lot in Example 3. Compare the widths of the confidence intervals.

a. Find t_c and E for each level of confidence.
b. Use $\bar{x}$ and E to find the left and right endpoints of each confidence interval.
c. Interpret the results and compare the widths of the confidence intervals.

Lesson 6-2 (continued) - Normal or t?

The flowchart describes when to use the standard normal distribution and when to use the t-distribution to construct a confidence interval for a population mean.



Notice in the flowchart that when both $n < 30$ and the population is *not* normally distributed, you *cannot* use the standard normal distribution or the t-distribution.

Example 4: Normal, t, or neither?

You randomly select 25 newly constructed houses. The sample mean construction cost is $181,000 and the population standard deviation is $28,000. Assuming construction costs are normally distributed, should you use the normal distribution, the t-distribution, or neither to construct a 95% CI for the mean construction cost?

Solution:

You randomly select 18 male athletes and measure the resting heart rate of each. The sample mean heart rate is 64 beats per minute with a sample standard deviation of 2.5 beats per minute. Assuming the heart rates are normally distributed, should you use the normal, t, or neither to construct a 90% CI for the mean heart rate?

Solution:

pg 330: 5, 6, 9, 13, 17, 21, 26 or 27, 30

Chapter 6: Confidence Intervals

$\hat{p}$

Lesson 6-3: CIs for Population Proportion

Population proportion p - tells us how many successes out of a fixed number of experiments

example: proportion of students at CHS that go to college

or, proportion of adults 18-25 who are Team Edward

Estimated using the sample proportion $\hat{p}$ pronounced "p hat"

DEFINITION

The **point estimate for *p***, the population proportion of successes, is given by the proportion of successes in a sample and is denoted by

$\hat{p} = \frac{x}{n}$ Sample proportion

where x is the number of successes in the sample and n is the sample size. The point estimate for the population proportion of failures is $\hat{q} = 1 - \hat{p}$. The symbols $\hat{p}$ and $\hat{q}$ are read as "p hat" and "q hat."

EXAMPLE 1

Finding a Point Estimate for *p*

In a survey of 1000 U.S. teens, 372 said that they own smartphones. Find a point estimate for the population proportion of U.S. teens who own smartphones.

Just like always, this true population parameter can be estimated, with a margin of error E - so we can make a confidence interval for the TRUE population proportion p using our error and p-hat

DEFINITION

A **c-confidence interval** for the population proportion p is

$$\hat{p} - E < p < \hat{p} + E$$

where

$$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

The probability that the confidence interval contains p is c.

The general process is very similar to previous CI construction - note that just like always with binomial, we need np and nq larger than 5

In Section 5.5, you learned that a binomial distribution can be approximated by a normal distribution when $np \geq 5$ and $nq \geq 5$. When $n\hat{p} \geq 5$, and $n\hat{q} \geq 5$, the sampling distribution of $\hat{p}$ is approximately normal with a mean of

$$u_{\hat{p}} = p$$

and a standard error of

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}.$$

$\left(\text{Notice } \sigma_{\hat{p}} = \frac{\sigma}{n} = \frac{\sqrt{npq}}{n} = \frac{\sqrt{npq}}{\sqrt{n^2}} = \sqrt{\frac{npq}{n^2}} = \sqrt{\frac{pq}{n}}.\right)$

Study Tip

Here are instructions for constructing a confidence interval for a population proportion on a TI-84 Plus.

STAT

Choose the TESTS menu.

A: 1–PropZInt . . .

Enter the values of *x*, *n*, and the level of confidence *c* (C-Level). Then select *Calculate*.

GUIDELINES

Constructing a Confidence Interval for a Population Proportion

IN WORDS	IN SYMBOLS
1. Identify the sample statistics n and x.	
2. Find the point estimate $\hat{p}$.	$\hat{p} = \frac{x}{n}$
3. Verify that the sampling distribution of $\hat{p}$ can be approximated by a normal distribution.	$n\hat{p} \geq 5$, $n\hat{q} \geq 5$
4. Find the critical value z_c that corresponds to the given level of confidence c.	Use Table 4 in Appendix B.
5. Find the margin of error E.	$E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}$
6. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\hat{p} - E$ Right endpoint: $\hat{p} + E$ Interval: $\hat{p} - E < p < \hat{p} + E$

EXAMPLE 2

Minitab and TI-84 Plus steps are shown on pages 344 and 345.

Constructing a Confidence Interval for *p*

Use the data in Example 1 to construct a 95% confidence interval for the population proportion of U.S. teens who own smartphones.

```
NORMAL FLOAT AUTO REAL DEGREE MP
1-PropZInt
x:372
n:1000
C-Level:0.95
Calculate
```

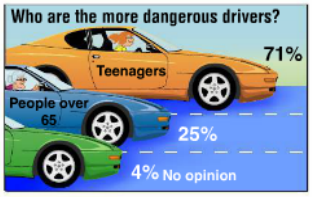
```
NORMAL FLOAT AUTO REAL DEGREE MP
1-PropZInt
(0.342,0.402)
p̂=0.372
n=1000
```

"We are ____% confident that the proportion of US teens who own smart phones is between ______ and ______."

Lesson 6-3 (continued) - Population proportion Confidence Intervals

Example 3: Constructing Confidence Intervals for *p*.

The graph shown is from a survey of 498 U.S. adults. Construct a 99% CI for the proportion of adults who think that teenagers are the more dangerous drivers.



Solution:

"We are ____% confident that the proportion who think teens are more dangerous is between _____ and ______."

As before, sometimes we want more precision with having to decrease our level of confidence - solution is still the same!

INCREASE SAMPLE SIZE!!!

FINDING A MINIMUM SAMPLE SIZE TO ESTIMATE *p*

Given a *c*-confidence level and a margin of error *E*, the minimum sample size *n* needed to estimate *p* is

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2.$$

This formula assumes that you have a preliminary estimate for $\hat{p}$ and $\hat{q}$. If not, use $\hat{p} = 0.5$ and $\hat{q} = 0.5$.

NOTE!!!

Example 4: Determining minimum sample size when using p-hat

You are running a political campaign and wish to estimate, with 95% confidence, the proportion of registered voters who will vote for your candidate. Your estimate must be accurate within 3% of the true population. Find the minimum sample size needed if 1) no preliminary estimate is available and 2) a preliminary estimate gives p-hat=0.31. Compare the results.

Solution:

Interpretation With no preliminary estimate, the minimum sample size should be at least 1068 registered voters. With a preliminary estimate of $\hat{p} = 0.31$, the sample size should be at least 914 registered voters. So, you will need a larger sample size when no preliminary estimate is available.

Try It Yourself 4

A researcher is estimating the population proportion of U.S. adults ages 18 to 24 who have had an HIV test. The estimate must be accurate within 2% of the population proportion with 90% confidence. Find the minimum sample size needed when (1) no preliminary estimate is available and (2) a previous survey found that 31% of U.S. adults ages 18 to 24 have had an HIV test.

a. Identify $\hat{p}$, $\hat{q}$, z_c, and E. If $\hat{p}$ is unknown, use 0.5.
b. Use $\hat{p}$, $\hat{q}$, z_c, and E to find the minimum sample size n.
c. Determine how many U.S. adults ages 18 to 24 should be included in the sample.

pg. 325
4,5,7,8,11,13,
17,18,23

Chapter 6: Confidence Intervals

$\chi_L{}^2$ $\chi_R{}^2$ σ

Lesson 6-4: CIs for Variation & Standard Deviation

At times, we want to control the amount something varies

ex: manufacturing auto parts to within certain tolerances so they are all very similar

DEFINITION

The **point estimate for σ^2** is s^2 and the **point estimate for σ** is s. The most unbiased estimate for σ^2 is s^2.

We have been using s as an approximation for σ - now we need a confidence interval for s to give us a range for the standard deviation (or, s^2 for σ^2 when we talk about variance).

*** When we talk about CIs for standard deviation or variance, we are using the **chi-squared distribution** ***

DEFINITION

If a random variable x has a normal distribution, then the distribution of

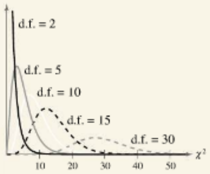
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

forms a **chi-square distribution** for samples of any size $n > 1$. Here are several properties of the chi-square distribution.

1. All values of χ^2 are greater than or equal to 0.
2. The chi-square distribution is a family of curves, each determined by the degrees of freedom. To form a confidence interval for σ^2, use the chi-square distribution with degrees of freedom equal to one less than the sample size.

   d.f. $= n - 1$ Degrees of freedom

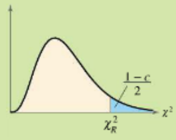
3. The total area under each chi-square distribution curve is equal to 1.
4. The chi-square distribution is positively skewed and therefore the distribution is not symmetric.
5. The chi-square distribution is different for each number of degrees of freedom, as shown in the figure. As the degrees of freedom increase, the chi-square distribution approaches a normal distribution.



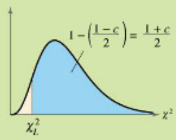
Chi-Square Distribution for Different Degrees of Freedom

Study Tip

For chi-square critical values with a c-confidence level, the values shown below, χ_L^2 and χ_R^2 are what you look up in Table 6 in Appendix B.

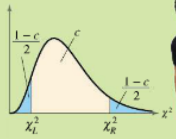


Area to the right of χ_R^2



Area to the right of χ_L^2

The result is that you can conclude that the area between the left and right critical values is c.



Finding chi-squared values requires two areas $\chi_L{}^2$ and $\chi_R{}^2$

Example 1: Finding chi-squared values

Find the critical values $\chi_L{}^2$ and $\chi_R{}^2$ for a 95% confidence interval when the sample size is 20.

Solution:

Degrees of freedom	α 0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170

χ_L^2 χ_R^2

From the table, you can see that $\chi_R^2 = 30.191$ and $\chi_L^2 = 7.564$.

Interpretation So, for a chi-square distribution curve with 17 degrees of freedom, 95% of the area under the curve lies between 7.564 and 30.191, as shown in the figure at the left.

Lesson 6-4: CIs for Variation & Standard Deviation

To find a confidence interval for the population standard deviation and variance:

DEFINITION

The *c*-confidence intervals for the population variance and standard deviation are shown.

Confidence Interval for σ^2:

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

Confidence Interval for σ:

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

The probability that the confidence intervals contain σ^2 or σ is c, assuming that the estimation process is repeated a large number of times.

GUIDELINES

Constructing a Confidence Interval for a Variance and Standard Deviation

IN WORDS	IN SYMBOLS
1. Verify that the population has a normal distribution.	
2. Identify the sample statistic n and the degrees of freedom.	d.f. $= n - 1$
3. Find the point estimate s^2.	$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$
4. Find the critical values χ_R^2 and χ_L^2 that correspond to the given level of confidence c and the degrees of freedom.	Use Table 6 in Appendix B.
5. Find the left and right endpoints and form the confidence interval for the population variance.	Left Endpoint, Right Endpoint: $\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$
6. Find the confidence interval for the population standard deviation by taking the square root of each endpoint.	Left Endpoint, Right Endpoint: $\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$

Note that the LEFT is using ${\chi_R}^2$ and the RIGHT is using ${\chi_L}^2$

Example 2: Constructing a CI for Variance and Standard deviation

You randomly select and weigh 30 samples of an allergy medicine. The sample standard deviation is 1.20 milligrams. Assuming the weights are normally distributed, construct 99% confidence intervals for the population variance and standard deviation.

Solution:

"We are ____% confident that the variance in weight is between _____ and ______, and that the standard deviation in weight is between _____ and _____."

Example 3: Constructing a CI for Variance and Standard deviation

The waiting time (in minutes) of a random sample of 22 people at a bank have a sample standard deviation of 3.6 minutes. Find a 98% confidence interval for the population variance and standard deviation.

Solution:

"We are ____% confident that the variance in wait time is between _____ and ______, and that the standard deviation in wait time is between _____ and _____."

pg. 334
3,4,7,9,10,15,19,21

Review Ch 6

CI for population mean μ
→ USE FLOWCHART!

* if $n \geq 30$ or if $n < 30$ but normal; we know σ (true pop st dev)

$$E = z_c \frac{\sigma}{\sqrt{n}} \qquad z_c : \left| \text{invnorm}\left(\frac{1-c}{2}\right) \right|$$

$$\bar{x} - E < \mu < \bar{x} + E$$

↑ sample mean

* if $n < 30$ and we don't σ, but we know s (sample st dev)

$$E = t_{df,c} \frac{s}{\sqrt{n}}$$

- df = n − 1
- use chart

$$\bar{x} - E < \mu < \bar{x} + E$$

* determining sample size needed to be within a given margin of error FOR MEAN

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

(estimate $\sigma \approx s$ if $n \geq 30$)

* CI for population proportion "p"
  - → Check $n\hat{p} \geq 5$, $n\hat{q} \geq 5$
    - $\hat{p}$ = sample proportion
    - $\hat{q} = 1 - \hat{p}$
  - → $E = z_c \sqrt{\frac{\hat{p}\hat{q}}{n}}$
  - → $\hat{p} - E < p < \hat{p} + E$
    - "proportion"
    - "82% of"

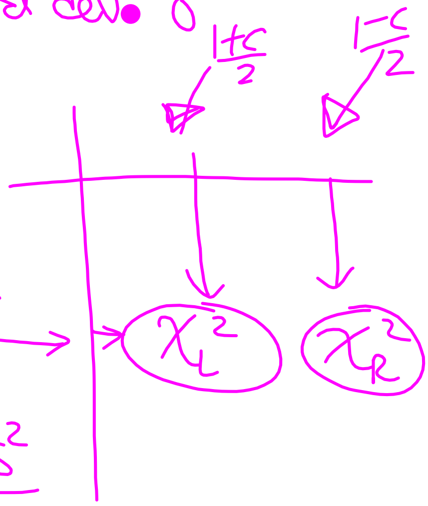
* determining sample size to be within a given margin of error: for proportions

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

if $\hat{p}$ is not given, use $\hat{p} = 0.5$, $\hat{q} = 0.5$

* CI for population variance σ^2 and " standard dev. σ

* use chi-squared table
  - df = n − 1
  - area R of $\chi_R^2 = \frac{1-c}{2}$
  - area R of $\chi_L^2 = \frac{1+c}{2}$



* $\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$

and $\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$

(s is sample st dev, s^2 is sample variance)

#1-3, p359