Ch. 7 – Statistical Intervals Based on a Single Sample

Before discussing the topics in Ch. 7, we need to cover one important concept from Ch. 6.

Standard error

The **standard error** is the standard deviation of the sampling distribution of a point estimate. It measures how much the point estimate or sample statistic varies from sample to sample. As n increases, the standard error decreases (and the point estimator tends to be closer to the population parameter it's estimating).

The **standard error of** $\bar{x}$ is equal to

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Example 1

A normal population has a mean of 50 and variance of 16. How large must the random sample be if we want the standard error of the sample average to be 1?

Based on the CLT, we know that the standard error is $\sigma_{\bar{x}} = \sigma/\sqrt{n}$. To solve this problem, we need to set the standard error equal to 1 and solve for n.

$$1 = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{16}}{\sqrt{n}} = \frac{4}{\sqrt{n}}$$

$$\frac{1}{4} = \frac{1}{\sqrt{n}}$$

$$4 = \sqrt{n}$$

$$n = 4^2 = 16$$

7.1 – Basic Properties of Confidence Intervals

Note from the Instructor: Please read pg. 268-272 (stopping after Example 7.3) in the book. I will not include the formulas from 7.1 in my notes because I am going to combine Formula 7.5 with Formula 7.8 from Section 7.2. I think this makes things a little bit more straight-forward and allows you to learn only one formula.

Confidence intervals (CIs) are a way to estimate population parameters, such as μ and p, based on a single sample selected from the population. CIs consist of a range around a point estimate within which the parameter is believed to fall, i.e. $\bar{x} \pm error$.

Example 2

In a random survey of 2500, 68% of Americans don't believe that they will receive social security at retirement (and thus must save for retirement on their own). The survey has a margin of error of ±3%.

a) What is the point estimate?

 68% is the point estimate because it's a single number calculated from the sample to estimate the population parameter we're interested in: the percentage of Americans that believe they will not get Social Security.

b) What is the interval estimate?

 The problem gives us a margin of error of ±3%. To get the interval estimate, we add and subtract the margin of error from the point estimate.

 68% ± 3% → (68 – 3, 68 + 3) → (65%, 71%)

Confidence intervals indicate the probable accuracy of the estimate, i.e. confidence level. The accuracy depends on the characteristics of the sampling distribution of the point estimate. The smaller the standard error of the estimate, the more accurate the estimate tends to be.

confidence level – the probability that the confidence interval contains the parameter (usually a number close to 1, such as 0.95 or 0.99, or a number close to 100% if expressed as a percentage)

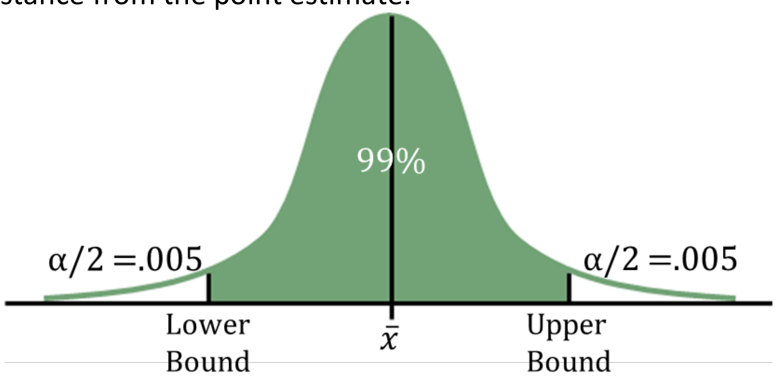
The “leftover” is called **α (alpha)** and is located in the two tails of the distribution. This is the probability that the CI does not contain the parameter.

$$\alpha = 1 - confidence\ level$$

Example 3

The National Center for Education Statistics surveyed 4400 college graduates about the length of time required to earn their bachelor’s degrees. The mean is 5.15 years and the standard deviation is 1.68 years. Construct a 99% confidence interval for the mean time required by all college graduates.

In this case, the confidence level is 99%. That means that $\alpha = 1 - .99 = .01$, which is split between the two tails of the normal distribution as shown below. This way the interval can be centered around the point estimate, $\bar{x}$, with the upper and lower bounds of the confidence interval equidistance from the point estimate.



Note: We will complete this example after defining the confidence interval formula.

7.2 – Large-Sample Confidence Intervals for a Population Mean and Proportion

A Large-Sample Interval for μ

The $100(1-\alpha)\%$ confidence interval for μ is given by

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

where $\bar{x}$ = sample mean, $z_{\alpha/2}$ = the z-value with an area of $\alpha/2$ to the right of it, σ = population standard deviation, and n = sample size.

Important Characteristics

1. If $n < 30$, sample must be from a Normal population and σ must be known.
2. If $n \geq 30$, s can be used if σ is unknown *(Note: The book requires it to be 40, but I will use 30 and expect you to do the same.)*
3. $Margin\ of\ error = m = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ *(Note: The book does not have a name for this value.)*
4. $width\ of\ interval = 2m$
5. To find the critical value, $z_{\alpha/2}$, you have to use the z-table with the negative values. Find the z that corresponds to the probability, $\alpha/2$, and then make the z-value positive. This works because the normal distribution is symmetric. I will also provide a list of commonly used critical values.
6. To decrease the width of the confidence interval, you must decrease the confidence level OR increase the sample size. The opposite is also true.

How to find $z_{\alpha/2}$

1. Find α using the confidence level
2. Calculate $\alpha/2$
3. Find the z that corresponds to the probability, $\alpha/2$, from Table A.3 on pg. A-6

Example 4

Find the critical value, $z_{\alpha/2}$, for the following confidence levels:

a) 97%

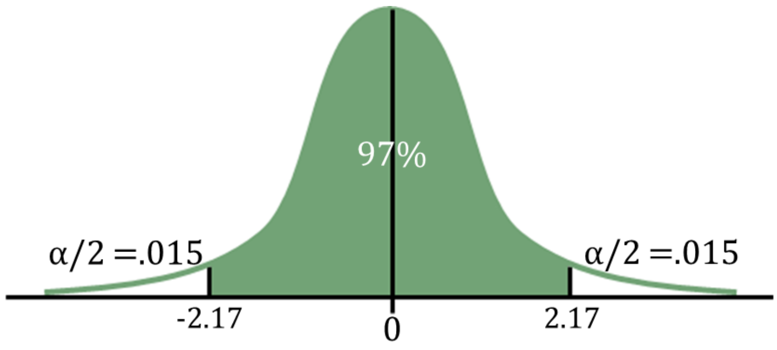
For a 97% CI, $\alpha = 1 - .97 = .03 \rightarrow \alpha/2 = .03/2 = .015$.

We need to find the z that corresponds to the probability .015. Let's look at the z table on page A-6 (reproduced below) because this gives us the probability in the lower tail of the 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	.0017	.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	.0038
−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	$\boxed{.0150}$	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183

Example 4 Continued

The red rectangle indicates the location of the probability .015 in the table. Now, you just move out to the edges of the table to determine the z. This probability corresponds to $P(Z < -2.17)$, but remember that a confidence interval is two-tailed like this:



The critical values are always positive because the negative is taken care of by the $\pm$ in the CI formula. Thus, $z_{.015} = 2.17$.

b) 95%

In this case, $\alpha = 1 - .95 = .05$, thus $\alpha/2 = .025$. Let's look for .025 in the table.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−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	.0352	.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

The red rectangle indicates where we found .025. Move your fingers out to the edge of the table to determine the value of z. Thus, $z_{.025} = 1.96$.

Commonly Used Critical Values, $z_{\alpha/2}$

Most of the time, you will be able to use the chart on the next page instead of the z table. You should include this on your formula sheet for exam 2.

Confidence Level	α	$\alpha/2$	$z_{\alpha/2}$
.90	.10	.05	1.645
.95	.05	.025	1.96
.96	.04	.02	2.05
.98	.02	.01	2.326
.99	.01	.005	2.576

Example 5 (Continuation of Example 3)

The National Center for Education Statistics surveyed 4400 college graduates about the length of time required to earn their bachelor's degrees. The mean is 5.15 years and the standard deviation is 1.68 years. Construct a 99% confidence interval for the mean time required by all college graduates.

We know 5.15 and 1.68 are the sample mean and standard deviation, respectively, because the first sentence mentions giving a survey to 4400 graduates. We can use the sample standard deviation in this case because the sample size is large.

$$n = 4400 \quad \bar{x} = 5.15 \quad s = 1.68 \quad \alpha = .01$$

Let's set up the confidence interval. We will use 2.576 as the critical value since this is a 99% confidence interval.

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 5.15 \pm 2.576\left(\frac{1.68}{\sqrt{4400}}\right) = 5.15 \pm 0.07 = (5.15 - 0.07, .5.15 + 0.07)$$

$$= (5.08, .5.22)$$

Finally, we must interpret the interval:
We are 99% confident that the mean time required for a bachelor's degree for all graduates is between 5.08 and 5.22 years.

Important note: You must ***always*** write a sentence interpreting the interval on quizzes and exams. This sentence should always include the confidence level, an explanation of the ***population*** parameter you're estimating, the bounds of the interval, and the units of the interval estimate. Notice that in the interpretation for this interval I say it's the mean time required for ALL graduates to distinguish the population.

Example 6

A 95% confidence interval for the average age of FSU students is (22, 28). Answer the following.

a) What was the value of $\bar{x}$ used in the problem?

It's easy to find this answer since the CI is symmetric. You just need to find the average of the upper and lower bounds.

$$\bar{x} = \frac{lower\ bound + upper\ bound}{2} = \frac{22+28}{2} = \frac{50}{2} = 25$$

b) What was the error used in the problem?

To find the sampling error, you can use the sample mean and the upper bound.

$$m = upper\ bound - \bar{x} = \ 28 - 25 = 3$$

Or you could calculate the range (or width) of the interval and divide by 2.

$$m = \frac{width}{2} = \frac{upper\ bound - lower\ bound}{2} = \frac{28-22}{2} = \frac{6}{2} = 3$$

c) President Stokes claims that the average age is 20. Is this reasonable? Explain.

No, because 20 does not fall within the 95% CI.

Determining the Sample Size Needed to Estimate μ

Note: You will not find this section in the book, but Example 7.7 on pg. 279 uses the following formula.

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

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

When finding n, you must *always* round up to the nearest whole number.

Example 7

An estimate is needed of the average age of customers in a department store correct to within 2 years with probability 0.98. How many customers should be sampled? Assume the standard deviation of the ages is 7.5.

We know that $\alpha = .02$, thus we will use 2.326 as the critical z-value. Since the store wants the estimate to be correct to within 2 years, we know that $m = 2$. We also know that $\sigma = 7.5$

$$n = \left(\frac{z_{\alpha/2}\sigma}{m}\right)^2 = \left(\frac{2.326 \cdot 7.5}{2}\right)^2 = \left(\frac{17.445}{2}\right)^2 = 8.7225^2 = 76.082 \approx 77$$

A Confidence Interval for a Population Proportion

Sometimes we might be interested in constructing an interval for the population proportion, p. This is essentially interval estimation for binomial data, where

$$proportion = \frac{\#\ of\ successes}{total\ sample\ size}$$

We are trying to estimate the parameter, $p =$ population proportion, with the statistic, $\hat{p} =$ sample proportion (pronounced p-hat).

Example 8

To estimate the proportion of American adults that are overweight, a random sample of 1000 adults is taken and 580 are classified as overweight. Identify p and $\hat{p}$.

$p =$ true proportion of overweight American adults
$\hat{p} =$ proportion of overweight in sample of 1000 $= 580/1000 =$ .58

The sampling distribution of $\hat{p}$ when n is large is approximately normal with a mean of p and standard deviation of $\sqrt{\frac{p(1-p)}{n}}$. Typically, to apply this approximation we require that $np \geq 5$ or $n(1-p) \geq 5$.

Large Sample $100(1-\alpha)\%$ CI for Population Proportion (p)

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

where $\hat{p}$ = sample proportion, $\hat{q} = 1 - \hat{p}$, $z_{\alpha/2}$ = the z-value with an area of $\alpha/2$ to the right of it, and n = sample size.

Example 9

According to a survey of 800 recent college graduates, 64% moved back home after graduation. Construct a 96% confidence interval for the true percentage of all recent college graduates that move back home after graduation.

$$n = 800 \qquad \hat{p} = .64 \qquad \alpha = 1 - .96 = .04 \qquad \alpha/2 = .02$$

Using the table of critical values of z provided earlier, we know that $z_{.02} = 2.05$. Now we can plug everything into the formula.

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .64 \pm 2.05\sqrt{\frac{.64(1-.64)}{800}} = .64 \pm 2.05(0.01697)$$

$$= .64 \pm .03479 \approx .64 \pm .04 = (.64 - .04, .64 + .04) = (0.60, 0.68)$$

Now, we can interpret the interval. We are 96% confident that the true percentage of all recent college graduates that move back home is between 60% and 68%.

Determining the Sample Size Needed to Estimate p

Note: You will not find this section in the book. Examples 7.8 and 7.9 in the book use a different method to find the sample size needed to estimate p. I want you to use the method that I describe below.

$$n = \left(\frac{z_{\alpha/2}}{m}\right)^2 p(1-p)$$

Important Note: An estimate of p is necessary to use the above formula. If an estimate is not provided, use $p = 0.5$.

Example 10

Suppose Monster.com is considering a poll to determine the percentage of college graduates that find a job before graduation. How many graduates must be surveyed to get a sample percentage with a margin of error of 1.5%? Use 95% confidence.

$m = .015 \quad \alpha = .05 \quad \alpha/2 = .025 \quad z_{.05} = 1.96$

Since p was not given, we will use 0.5.

$$n = \left(\frac{z_{\alpha/2}}{m}\right)^2 p(1-p) = \left(\frac{1.96}{.015}\right)^2 (.5)(1-.5) = 17073.8(.5)(.5) = 4268.44 \approx 4269$$

One-Sided Confidence Intervals (Confidence Bounds)

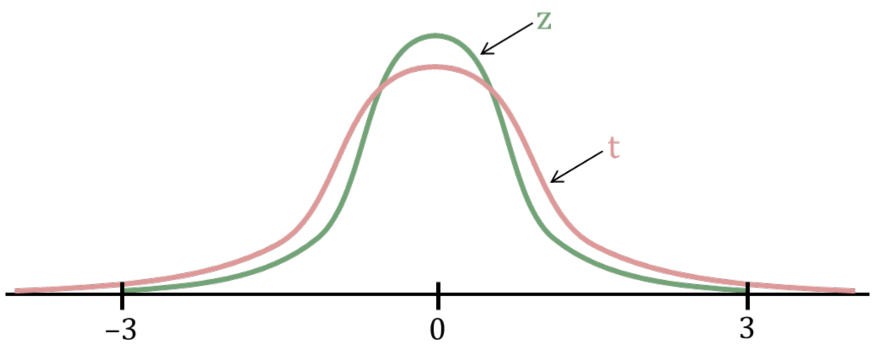
Please skip this section.

7.3 – Intervals Based on a Normal Population Distribution

When the population variance is known or the sample size is large ($n \geq 30$), you can use the confidence interval for μ that was provided in section 7.2. However, when the sample size is small ($n < 30$), we must use a different technique because x may not be normal (violates CLT) and s is not a good estimate of σ. For small sample confidence intervals, we must use the t distribution and assume a normal population.

Properties of t distributions

1. Mound-shaped and symmetric around 0
2. More variable than z



3. Many different t distributions exist since t depends on degrees of freedom ($v = n - 1$)

4. As $n \to \infty$, $t \to z$ (that is why the numbers at the bottom of the t table match the critical z-values)
5. $t_{\alpha,v}$ = the value of t that has α of the t_v distribution to the right of it. This value is called a ***t* critical value**. The t table on pg. A-9 gives t critical values for specified values of α and v (Note: This table does NOT contain probabilities like the Z tables.)

How to Use the t Table

1. Determine the value of α (usually given in the problem)
2. Calculate the degrees of freedom, $v = n - 1$
3. Find the correct value of α at the top of the table and place right finger there
4. Find the correct value of v on the left side of the table and place left finger
5. Simultaneously, slide right finger down the column and left finger across the row until your fingers meet at $t_{\alpha,v}$

Example 11

Find the t critical values for the following values of α and n.

a) $\alpha = .10, n = 2$

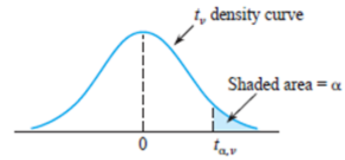
A reproduction of the t table is on the next page. It gives critical values of t for specified values of α and degrees of freedom, v. The values of α are across the top labeling the table's columns. The values of v are down the side labeling the rows. For this problem, we are only interested in the $\alpha = .10$ column (boxed in red).

(Example continued on next page.)

Example 11 Continued

Appendix Tables A-9

Table A.5 Critical Values for t Distributions



	α						
v	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Now we need to calculate the degrees of freedom to find the correct row.

$$v = n - 1 = 2 - 1 = 1$$

Thus, we are interested in the first row of the table (as seen on the next page). The value of t we're looking for is where the 2 red boxes intersect.

Example 11 Continued

	α						
ν	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610

Thus, $t_{.10,1} = 3.078$.

b) $\alpha = .05, n = 21$

We know that we are interested in the column for $\alpha = .05$. Let's determine what row we need to look at.

$$v = n - 1 = 21 - 1 = 20$$

The value we're looking for should be in the 20th row and the $\alpha = .05$ column.

	α						
ν	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792

Thus, $t_{.05,20} = 1.725$.

The One-Sample t Confidence Interval

When a small sample is taken from a normal population and σ is unknown, you should use the one-sample t confidence interval or what I like to call the small sample $100(1-\alpha)\%$ confidence interval for μ.

Small Sample 100(1 – α)% CI for μ ($n < 30$)

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

where $\bar{x}$ = sample mean, $t_{\alpha/2,n-1}$ = the t value with an area of $\alpha/2$ to the right of it based on $n-1$ degrees of freedom, s = sample standard deviation, and n = sample size.

Example 12

16 loan applications were randomly selected and the dollar amount requested has a mean of \$900 and a standard deviation of \$120. Construct and interpret a 95% confidence interval for mean dollar amount of all loans requested at the bank.

Below is all the information given in the problem:
$n = 16$ $\quad \bar{x} = 900$ $\quad s = 120$ $\quad \alpha = 1 - .95 = .05$

Since the sample size is small ($n < 30$), we will need to use the t table. First, we have to calculate the degrees of freedom.
$v = n - 1 = 16 - 1 = 15$

So, we know we will be looking at the 15th row of the t table. But what column do we need to look at? Remember that α gets split between the two tails for a confidence interval. Thus, we need to find $\alpha/2$ and that will be the column we look at in the t table.
$\frac{\alpha}{2} = \frac{.05}{2} = .025$

The t value in the 15th row and the .025 column is $t_{.025,15} = 2.131$. Now we can plug all the values into the small sample CI formula.

$$\bar{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}} = 900 \pm 2.131\left(\frac{120}{\sqrt{16}}\right) = 900 \pm 2.131(30) = 900 \pm 63.93$$

$$= (900 - 63.93, .900 + 63.93) = (836.07, .963.93)$$

We are 95% confident that the mean dollar amount of all loans requested at the bank is between \$836.07 and \$963.93.

Example 13

Follow-up question to Example 12: If the bank manager said the mean was $1000, would you believe him based on the CI?

No, because $1000 is not in the CI.

Example 14

Suppose you want to estimate the average age of FSU students. A sample of 5 students is selected and the following ages are recorded: 20, 18, 19, 28, and 18. Find a 99% confidence interval for the average age of all FSU students. Interpret.

For this example, the sample mean and standard deviation is not provided, so we will have to use what we learned in Ch. 6 to calculate them.

$$\bar{x} = \frac{\sum x}{n} = \frac{20 + 18 + 19 + 28 + 18}{5} = \frac{103}{5} = 20.6$$

I will leave the calculation of the standard deviation for you to do on your own. You should get $s = 4.22$.

Now we need to find the critical t value. Since the confidence level is 99%, we know that $\alpha = 1 - .99 = .01$, thus $\alpha/2 = .01/2 = .005$. Since $n = 5$, we have 4 degrees of freedom. So, we need to find the t in the .005 column and the 4th row.

$$t_{.005,4} = 4.604$$

Now, we can construct the confidence interval.

$$\bar{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}} = 20.6 \pm 4.604\left(\frac{4.22}{\sqrt{5}}\right) = 20.6 \pm 4.604(1.887) = 20.6 \pm 8.69$$

$$= (20.6 - 8.69, .20.6 + 8.69) = (11.91, .29.29)$$

We are 99% confident that the average age of all FSU students is between 11.91 and 29.29 years.

A Prediction Interval for a Single Future Value

Please skip this section.

Tolerance Intervals

Please skip this section.

Intervals Based on Nonnormal Population Distributions

Please skip this section.

7.4 – Confidence Intervals for the Variance and Standard Deviation of a Normal Population

Although many practical problems involve inferences about a population mean, it is sometimes of interest to make an inference about a population variance or standard deviation. To illustrate, a quality control supervisor in a cannery knows that the exact amount each can contains will vary, since there are certain uncontrollable factors that affect the amount of fill. If σ^2 is large, some cans will contain too little and others too much. Suppose regulatory agencies specify that the standard deviation of the amount of fill in 16-ounce cans should be less than .1 ounce. The quality control supervisor needs a way to determine if the cannery's process is meeting this specification. One way to do this is to sample some cans and construct a confidence interval for the population variance. CIs for the variance require the use of the chi-squared distribution.

Chi-Squared (χ^2) Distribution

Let $X_1, X_2, \dots, X_n$ 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

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

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

Characteristics of the χ^2 distribution

1. Values are always positive
2. Not symmetric, skewed to the right
3. Mean and variance depend on degrees of freedom and thus on n
4. As $n \to \infty$, the distribution becomes more symmetric
5. χ^2 table on pg. A-11 gives critical values (or percentage points) of χ^2 for specified values of α and degrees of freedom (v) (similar to t table)

How to use the χ^2 Table

1. Determine the value of α (usually given in the problem)
2. Calculate the degrees of freedom, $v = n - 1$
3. Find the correct value of α at the top of the table and place right finger there
4. Find the correct value of v on the left side of the table and place left finger
5. Simultaneously, slide right finger down the column and left finger across the row until your fingers meet at $\chi^2_{\alpha,v}$

Example 15

Find the χ^2 critical values for the following values of α and n.

a) $\alpha = .1, n = 8$

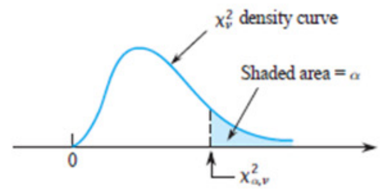
On the next page, is a reproduction of the χ^2 table. It gives critical values of χ^2 for specified values of α and degrees of freedom, v. The values of α are across the top labeling the table's columns. The values of v are down the side labeling the rows. For this problem, we are only interested in the $\alpha = .10$ column (boxed in red).

(Example continued on the next page.)

Example 15 Continued

Appendix Tables **A-11**

Table A.7 Critical Values for Chi-Squared Distributions



	α									
ν	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
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	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
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	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

Next, we need to determine what row we should look in. We do this by calculating the degrees of freedom, ν.

$$\nu = n - 1 = 8 - 1 = 7$$

Example 15 Continued

Thus, we are interested in the 7th row (boxed in red on the next page). The value of χ^2 we're looking for is where the two red rectangles intersect.

ν	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
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	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

Thus, $\chi^2_{.1,7} = 12.017$.

b) $\alpha = .9, n = 25$

We know that we are interested in the column for $\alpha = .90$. Let's determine what row we need to look at.

$$v = n - 1 = 25 - 1 = 24$$

The value we're looking for should be in the 24th row and the $\alpha = .90$ column.

ν	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925

Thus, $\chi^2_{.9,24} = 15.659$.

100(1 – α)% Confidence Interval for σ²

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 $100(1-\alpha)\%$ confidence interval on σ^2 is given by

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$$

where $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-\alpha/2,n-1}$ are the lower and upper $\alpha/2$ critical values of the chi-square distribution with $n-1$ degrees of freedom, respectively. A **confidence interval for the standard deviation σ** has lower and upper bounds that are the square roots of the corresponding limits above.

Example 16

Suppose that the quality control supervisor at the cannery (mentioned earlier) randomly selects 10 cans and weighs the contents of each. He finds that the sample standard deviation is 0.0412. Construct a 90% confidence interval and determine if the canning process is violating the regulation that the standard deviation of the can weights should be less than .1 ounce.

In this case, $n = 10$, $s = .0412$, $\alpha = 1 - .90 = .10$, and $\alpha/2 = .05$. The degrees of freedom for the χ^2 critical values is $v = 10 - 1 = 9$. Therefore, we need to find the χ^2 value for 9 degrees of freedom with $\alpha/2 = .05$ and $1 - \alpha/2 = .95$.

$$\chi^2_{.05,9} = 16.919 \qquad \chi^2_{.95,9} = 3.325$$

Now, we can plug all the values into the CI formula. First, let's calculate the lower bound.

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} = \frac{(10-1)(0.0412)^2}{16.919} = \frac{9(0.001697)}{16.919} = \frac{0.015273}{16.919} = 0.0009$$

Then, calculate the upper bound.

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} = \frac{0.015273}{3.325} = 0.0046$$

Therefore, the 90% interval for the variance is $(0.0009, 0.0046)$. In order to properly interpret it, we should find the interval for the standard deviation since it is in the same units as $X = can\ weights$. To find the interval for the standard deviation, we just take the square root.

$$\left(\sqrt{0.0009}, \sqrt{0.0046}\right) = (0.03, 0.0678)$$

Example 16 Continued

We are 90% confident that the standard deviation of all the can weights is between 0.03 and 0.0678 ounces.

Now, we need to answer the question: is the cannery violating the regulatory agencies' requirement that the standard deviation be less than .1 ounce? No, because the upper and lower bounds of the 90% confidence interval are less than .1 ounce.