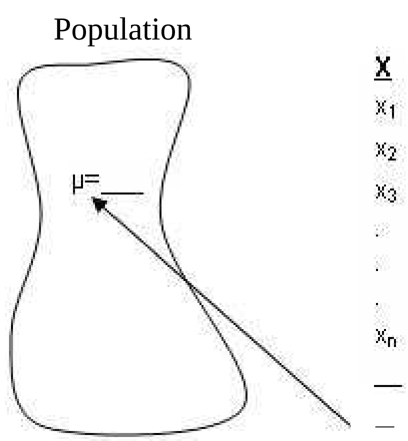
Chapter 9—Confidence Intervals about a Single Parameter

9.1 Confidence Intervals about a Population Mean

Point Estimate vs. Interval Estimate

Definition: A **point estimate** of a parameter is the value of a statistic ($\bar{X}$) that estimates the value of the parameter (μ).



$\bar{x}$ =39 -- **point estimate**. Will $\bar{x}$ be a perfect estimate of μ?

|-------- $\bar{x}$ -------| -- **interval estimate** of μ.

36 ↑ **lower bound** 42 ↑ **upper bound**

Definition: A confidence interval estimate of a parameter consists of an interval of numbers, along with a measure of the likelihood that the interval contains the unknown parameter. The level of confidence in a confidence interval is the proportion of the intervals that will contain μ if a large number of repeated samples are obtained. The level of confidence is denoted (1-α)·100%.

Note that a confidence interval includes:

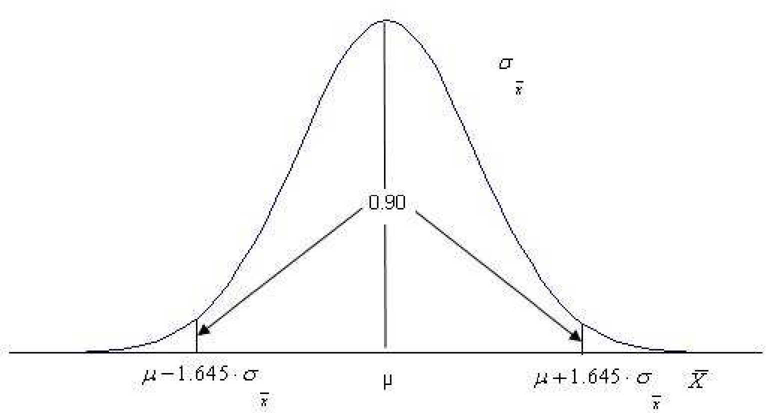
- a lower bound
- an upper bound
- an interval equal to the upper bound minus the lower bound
- a level of confidence that the interval contains μ

Chapter 9 covers three types of confidence intervals:

- 9.1—Confidence interval for μ with σ known.
- 9.2—Confidence interval for μ with σ unknown.
- 9.3—Confidence interval for the population proportion, p.

Derivation of a 90% Confidence Interval (CI):

Sampling Distribution of $\bar{X}$



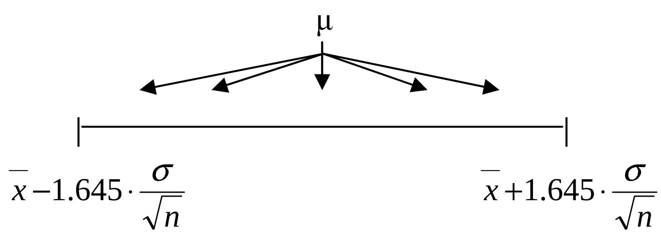
Because $\bar{X}$ is normally distributed, we know that 90% of all sample means should lie within 1.645 standard deviations of the population mean, μ, i.e.,

$$\mu - 1.645 \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1.645 \cdot \frac{\sigma}{\sqrt{n}}$$

With a little algebraic manipulation (adding -μ-$\bar{X}$ across the equation and then multiplying by -1), we can rewrite this equation with μ in the middle:

lower bound   upper bound

$$\bar{X} - 1.645 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.645 \cdot \frac{\sigma}{\sqrt{n}}$$



A succinct way to write a 90% CI is: $\bar{X} \pm 1.645 \cdot \frac{\sigma}{\sqrt{n}}$

↑ point estimate ($\bar{X}$)   ↑ margin of error ($1.645 \cdot \frac{\sigma}{\sqrt{n}}$)

> Confidence interval:
> Point estimate ±
> Margin of error.

Constructing a (1-α)·100% Confidence Interval about μ, σ Known (p. 344)

(1-α) ·100% CI: $\bar{x} \pm E(margin\ of\ error)$

where $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ and $z_{\alpha/2}$ is the critical Z-value.

Note: The size, n, of the sample must be greater than or equal to 30 or the population must be normally distributed (see section 9.1).

Example: 90% CI for mean price of 3-year old Corvette

Table 1--Price of 3-Year Old Corvette ($ per car)

47,000	43,108	33,995
32,750	33,988	43,500
33,995	32,750	39,950
36,900	35,995	39,998
37,995	37,995	43,785

$\bar{X} = 38{,}247 \quad \sigma = 4{,}100 \quad n = 15 \quad 90\%\ confidence \rightarrow \alpha = 0.10$

Formula—90% CI: $\bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$\bar{X} \pm 1.645 \cdot \frac{\sigma}{\sqrt{n}}$$

<u>lower bound</u>

$38{,}247 - 1.645 \cdot \frac{4{,}100}{\sqrt{15}}$

$38{,}247 - 1{,}741$

$36{,}506$

<u>upper bound</u>

$38{,}247 + 1.645 \cdot \frac{4{,}100}{\sqrt{15}}$

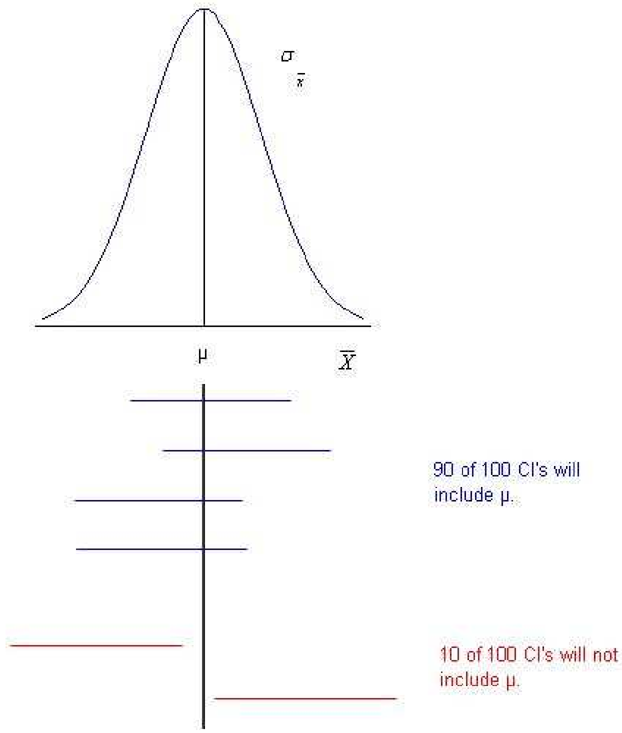
$38{,}247 + 1{,}741$

$39{,}988$

Interpretation of a Confidence Interval (p. 410)

A $(1-\alpha)\cdot 100\%$ confidence interval means that if we obtained many simple random samples of size n from the population whose mean, μ, is unknown, then approximately $(1-\alpha)\cdot 100\%$ of the intervals will contain μ.

For example, assume we draw 100 random samples of size n=15 and compute one hundred 90% CI's.



Interpretation of the 90% CI for 3-year-old Corvettes (derived on p. 3):

- We are 90% confident that the interval \$36,506 to \$39,988 will contain the unknown population mean price, μ.
- If we were to draw 100 random samples and compute 100 CI's, 90 of the CI's would contain μ.

Will the 90% CI (for the mean price of 3-year old Corvettes) derived on p. 3 include the unknown population mean μ? We can't say for sure, but using the concept of *confidence* described above, we are 90% sure that the interval will contain the true population mean.

The Role of the Margin of Error (p. 412)

> *Definition*: **The Margin of Error**
>
> The margin of error, E, in a (1-α)·100% confidence interval in which σ is known is given by
>
> $$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
>
> where n is the sample size.
> **Note:** We require the population from which the sample was drawn be normally distributed or the sample size n be greater than or equal to 30.

|--|

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

|◄——— E ———►|◄——— E ———►|

Note: The total length of the confidence interval = 2E.
One-half of the confidence interval = E.

How can we reduce the width of the confidence interval? The width of the interval is determined by the margin of error:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

The margin of error depends on 3 quantities:

(1) Level of confidence, (1-α)·100%
(2) Standard deviation of the population, σ
(3) Sample size, n

A statistician cannot affect (2) but he can control (1) and (3).

How can you reduce the size of the confidence interval? Use the Corvette example from p. 3.

$$E = z_{0.10/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.645 \cdot \frac{4{,}100}{\sqrt{15}} = 1{,}741$$

$90\%\ CI: \quad \bar{x} \pm margin\ of\ error\ (E)$

$(n = 15) \quad 38{,}247 \pm 1{,}741$

$36{,}506\ to\ 39{,}988$

<u>Increase sample size, n:</u>
n increases from 15 to 25.

$$E = z_{0.10/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.645 \cdot \frac{4{,}100}{\sqrt{25}} = 1{,}349$$

$90\%\ CI: \quad \bar{x} \pm margin\ of\ error\ (E)$

$(n = 25) \quad 38{,}247 \pm 1{,}349$

$36{,}898\ to\ 39{,}596$

Summary: As the sample size *increases*, the size of the margin of error and the confidence interval *decreases*.

<u>Reduce confidence level:</u>.
Confidence level is decreased from 90 to 75.

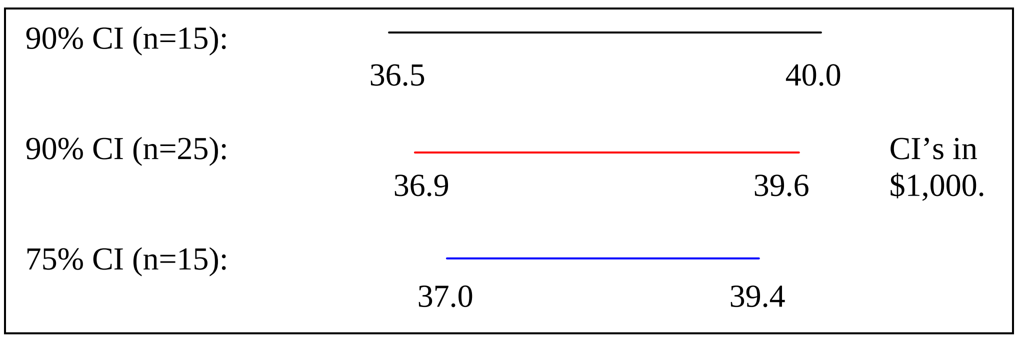
$$E = z_{0.25/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.15 \cdot \frac{4{,}100}{\sqrt{15}} = 1{,}217$$

$75\%\ CI: \quad \bar{x} \pm margin\ of\ error\ (E)$

$(n = 15) \quad 38{,}247 \pm 1{,}217$

$37{,}030\ to\ 39{,}464$

Summary: As the level of confidence *decreases*, the size of the margin of error and the confidence interval *decreases*.



Work problems #15, p. 350 and #11, p. 349.

Determining the Sample Size

Determining the Sample Size n (p. 413)

The sample size required to estimate the population mean, μ, with a specified margin of error, E, is given by

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

where n is *rounded up* to the nearest whole number.

The formula for n is derived from the margin of error formula:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

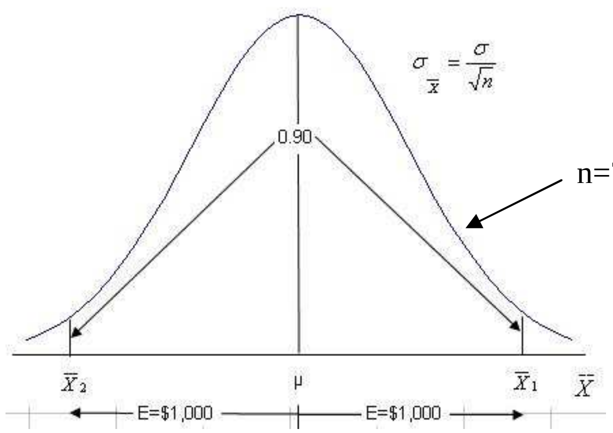
See p. 347.

Confidence level → Sample size, n ← Margin of error

Example: Consider estimating the mean price of a 3-year old Corvette. How large a sample is required to estimate the mean price within $1,000 with 90% confidence?

- 90% confidence >>> α=0.10 >>> $z_{0.10/2} = z_{0.05} = 1.645$
- Population standard deviation is assumed known at σ=$4,100
- Margin of error is E=$1,000

Sampling Distribution of $\overline{X}$



Substitute the values of z, σ, and E into the formula for determining the sample size and obtain

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 = \left(\frac{1.645 \cdot 4{,}100}{1{,}000} \right)^2 = 6.7445^2 = 45.488 \text{ or } 46$$

Work problem #27, p. 354.

9.2 Confidence Intervals about a Population Mean, σ Unknown

In Section 9.1, we made the assumption that the population standard deviation, σ, was *known*. In Section 9.2, we will drop this assumption and assume that σ is *unknown*.

The CI that we derived on p. 2 was based on the z-distribution, i.e., $z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$. The problem with using the z-distribution is that it requires knowledge of the population standard deviation, σ. But, it is difficult to know a population parameter like σ, which is typically unknown to a practicing statistician.

Remember that σ is a *parameter* and s is a *statistic*. Statistics are used to estimate parameters.

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} \text{ and } \sigma = \sqrt{\sigma^2}$$

↑ estimates

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} \text{ and } s = \sqrt{s^2}$$

W.S. Gossett first studied the problem of using s to estimate the unknown σ.

- Gossett was in charge of conducting experiments at the Guinness Brewery to identify the best barley varieties.
- He performed experiments with small data sets and σ was not known.
- At that time, the only available distribution was the standard normal distribution,

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}.$$

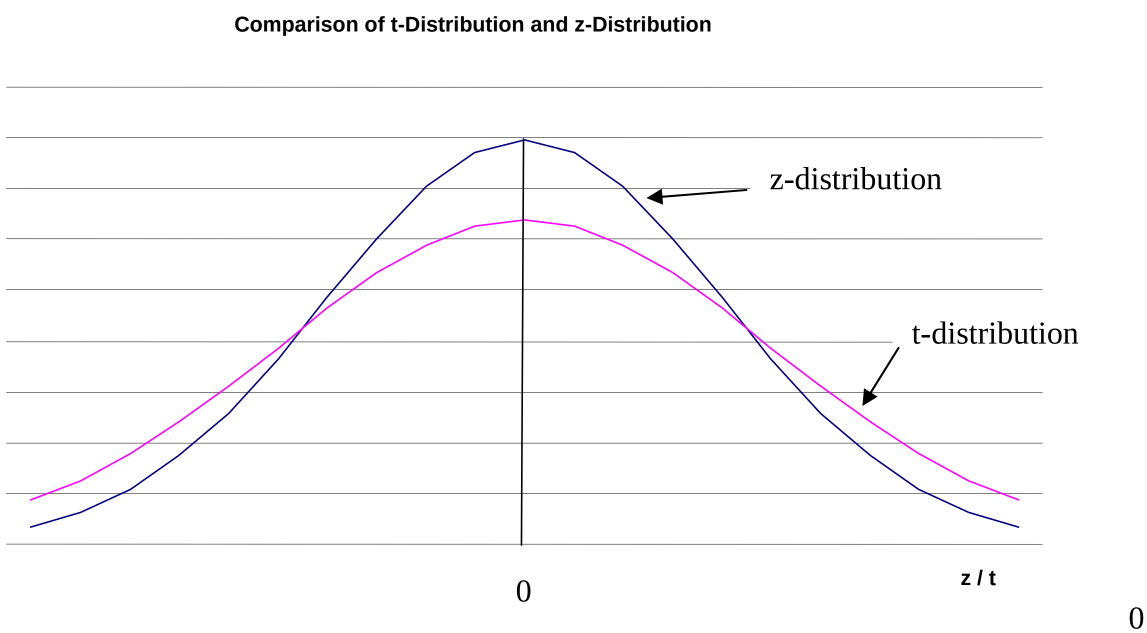
Gossett replaced σ with s and developed the sampling distribution of $\frac{\bar{X} - \mu}{s / \sqrt{n}}$. The distribution Gossett developed is not a normal distribution but rather followed a new distribution called Student's t-distribution (Gossett published his findings under the pen name of "Student").

Student's t-Distribution

Suppose a simple random sample of size n is taken from a population. If the population from which the sample is drawn follows a normal distribution, then the distribution of

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

follows Student's t-distribution with n-1 degrees of freedom (df), where $\bar{x}$ is the sample mean and s is the sample standard deviation.



Properties of the t-Distribution, p. 425.

1. Centered at the mean of μ=0.
2. Bell-shaped curve that is symmetric about 0. The area under the curve to the right of 0 equals the area under the curve to the left of 0 equals ½.
3. Total area under the curve is 1.
4. Asymptotic to the horizontal axis (i.e., as t increases or decreases without bound, the graph approaches the horizontal axis).
5. Area in the tails of the t-distribution is a little greater than the area in the tails of the standard normal distribution, because we are using "s" as an estimate of σ, thereby introducing further variability into the t-statistic as compared to the z-statistic.

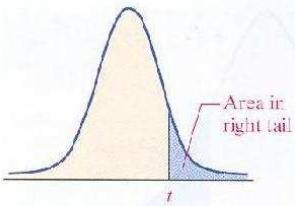


TABLE III

t-Distribution

Area in Right Tail

df=n-1	0.25	0.20	0.15	0.10	0.05	0.025	0.02	0.01	0.005	0.0025	0.001	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.706	15.894	31.821	63.657	127.321	318.289	636.558
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.089	22.328	31.600
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.214	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
31	0.682	0.853	1.054	1.309	1.696	2.040	2.144	2.453	2.744	3.022	3.375	3.633
32	0.682	0.853	1.054	1.309	1.694	2.037	2.141	2.449	2.738	3.015	3.365	3.622
33	0.682	0.853	1.053	1.308	1.692	2.035	2.138	2.445	2.733	3.008	3.356	3.611
34	0.682	0.852	1.052	1.307	1.691	2.032	2.136	2.441	2.728	3.002	3.348	3.601
35	0.682	0.852	1.052	1.306	1.690	2.030	2.133	2.438	2.724	2.996	3.340	3.591
36	0.681	0.852	1.052	1.306	1.688	2.028	2.131	2.435	2.719	2.990	3.333	3.582
37	0.681	0.851	1.051	1.305	1.687	2.026	2.129	2.431	2.715	2.985	3.326	3.574
38	0.681	0.851	1.051	1.304	1.686	2.024	2.127	2.429	2.712	2.980	3.319	3.566
39	0.681	0.851	1.050	1.304	1.685	2.023	2.125	2.426	2.708	2.976	3.313	3.558
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
70	0.678	0.847	1.044	1.294	1.667	1.994	2.093	2.381	2.648	2.899	3.211	3.435
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
90	0.677	0.846	1.042	1.291	1.662	1.987	2.084	2.368	2.632	2.878	3.183	3.402
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	0.25	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**

Area in Right Tail

Definition: The **number of degrees of freedom (df)** is the maximum number of variates that can freely be assigned (i.e., calculated) before the rest of the variates are completely determined. Or, stated in another way, the number of df is the number of deviations that must be known before the remaining deviation(s) can be calculated.

Example of calculation of df (x_1=4, x_2=5, x_3=9).

	X		$(X_i - \overline{X})$
	4		4-6 = -2
	5		5-6 = -1
	9		
$\overline{X}$ =	**6**	$\sum(X_i - \overline{X})$	= 0

Finding t-Values in Table III, p. A-4:

#1. For sample size n=15 and α=0.10, find $t_{0.10}$. Note that df=15-1=14.

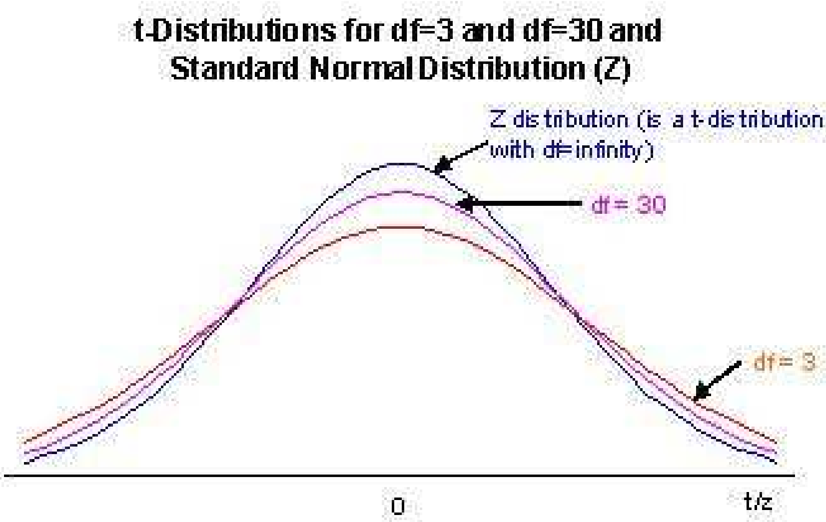
#2. For sample size n=15 and α=0.10, find $\pm t_{0.10/2} = \pm t_{0.05}$ (i.e., the upper and lower t-values).

#3. For sample size n=43 and α=0.05, find $-t_{0.05}$. Note if the df we desire is not available in Table III, we will follow the practice of choosing the closest df (in this case, df=40).

#4. For sample size n = ∞ and α=0.05, find $t_{0.05}$. Notice that $t_{0.05}$ is the same as $z_{0.05}$. Why does this occurs?

Additional Properties of the t-Distribution, p. 425

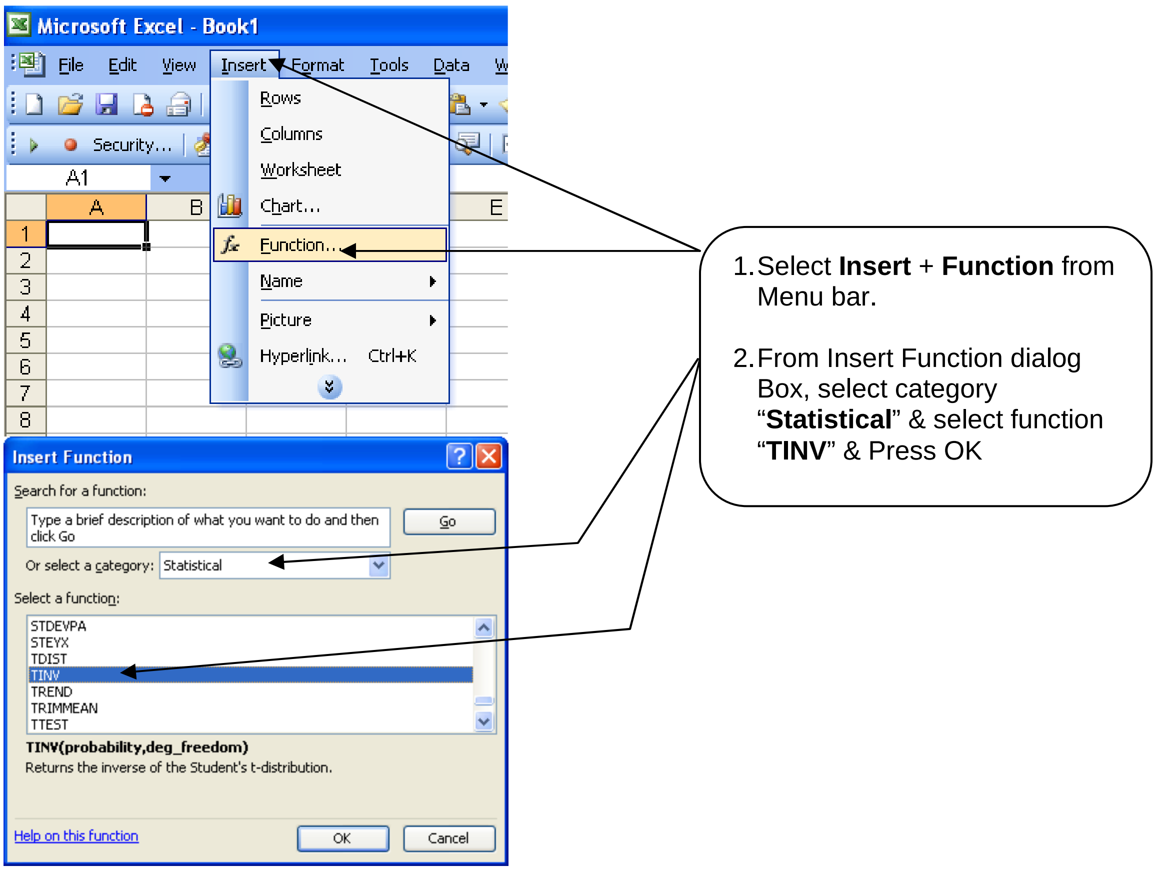
1. As the sample size n increases, the t-curve gets closer to the standard normal curve. This result occurs because, as the sample size n increases, the values of "s" get closer to the value of σ (i.e., the dispersion of the statistic s about σ decreases).
2. The t-distribution is different for different values of n, the sample size.

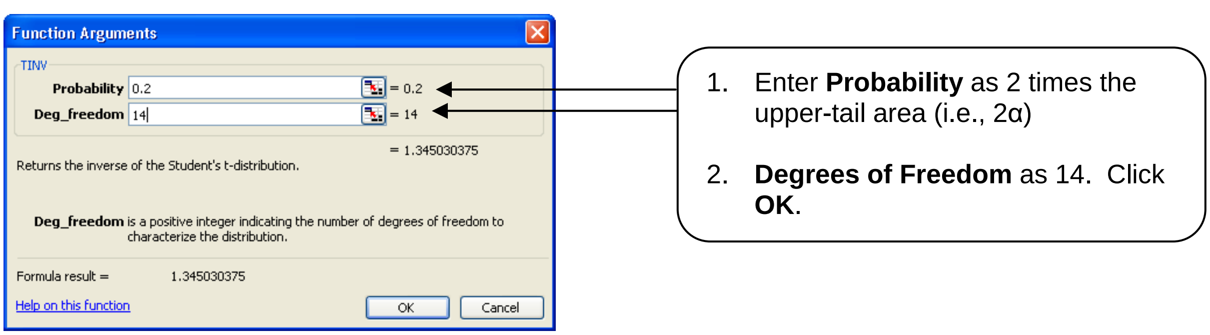


Finding t-Values Using Excel:

#1. For sample size n=15 and α=0.10, find $t_{0.10}$. Note that df=15-1=14.

Excel: Finding t-values for the upper tail of the t-distribution:
Step 1: Select **Insert/Function** (***f***$_x$) from the Windows menubar. In the **Function Category**, select "**Statistical**." In the **Function Name**, select "**TINV**." Click **OK**.
Step2: Enter probability as 2 times the upper-tail area (i.e., 2α) and degrees of freedom as 14. Click **OK**.



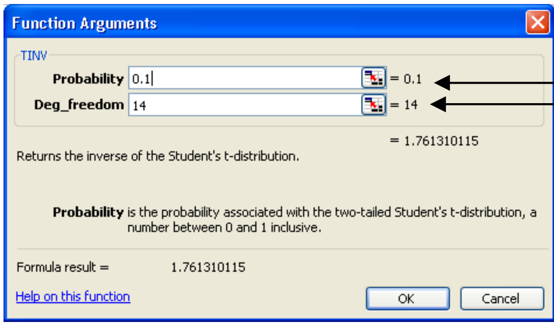


#2. For sample size n=15 and α=0.10, find $\pm t_{0.10/2} = \pm t_{0.05}$ (i.e., the upper and lower t-values).

Excel: Finding t-values for the upper and lower tails of the t-distribution:

Step 1: Select **Insert/Function** (***f**$_x$*) from the Windows menubar. In the **Function Category**, select "Statistical." In the **Function Name**, select "TINV."

Step2: Enter probability as α=0.10 and degrees of freedom as 14. Click **OK**.



1. Follow initial steps mentioned in **Ex. #1**
2. Enter **Probability** as as α=0.10
3. **Degrees of Freedom** as 14. Click **OK**.

Constructing a Confidence Interval for μ, σ Unknown.

$$(1-\alpha)\cdot 100\%\ CI: \quad \bar{X} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Example—Consider developing a 90% confidence interval for the mean price of a 3-year old Corvette.

Table 1--Price of 3-Year Old Corvette ($ per car)

47,000	43,108	33,995
32,750	33,988	43,500
33,995	32,750	39,950
36,900	35,995	39,998
37,995	37,995	43,785

$$\bar{X} = 38{,}247 \qquad n = 15 \qquad 90\%\ confidence >>> \alpha = 0.10$$

$$90\%\ CI: \quad \bar{X} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$= 38{,}247 \pm 1.761 \cdot \frac{4{,}522}{\sqrt{15}}$$

$$= 38{,}247 \pm 2{,}056$$

$$= 36{,}191\ to\ 40{,}303$$

The sample standard deviation, s, is calculated using the formula:

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} = \frac{\sum X_i^2 - n\bar{X}^2}{n-1}$$

Points to note about the t confidence interval procedure:

- When the population is normally distributed, the procedure is exact.
- For moderate departures from the normally assumption, the procedure is accurate.
- For large n, the procedure is approximately correct for non-normal populations.
- If the data set is small and there are outliers, the procedure is compromised.

Work problems #1 and #3, p. 364.

9.3 Confidence Interval for a Population Proportion

Probably the most frequently reported confidence interval is one involving the proportion of a population. In a March 2005 *FOX News/Opinion Dynamics Poll*, a random sample of n=900 registered voters nationwide was asked "Do you approve or disapprove of the job George W. Bush is doing as president?" The approval rating on 03/1-2/2005 was 52% with a margin of error of ± 3% at 95% confidence (http://www.pollingreport.com/BushJOb.htm).

Point Estimate of a Population Proportion

Suppose simple random sample of size n is obtained from a population in which each individual either does or does not have a certain characteristic. The best point estimate of p, denoted $\hat{p}$, (read "p-hat"), the proportion of the population with a certain characteristic, is given by

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

where x is the number of individuals in the sample with the specified characteristic.

Sampling Distribution of $\hat{p}$

For a simple random sample of size n such that n $\leq$ 0.05N (that is, the sample size is no more than 5% of the population size), the sampling distribution of $\hat{p}$ is approximately normal with mean $\mu_{\hat{p}} = p$ and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$, provided that np(1-p) $\geq$ 10.

Constructing a (1-α)·100% Confidence Interval for a Population Proportion

(1-α)·100% CI: point estimate ± margin of error

$$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{(\hat{p}(1-\hat{p})}{n}}$$

Check the assumptions:

1. n≤0.05N
2. np(1-p)≥10

Example—95% CI for President Bush's Job Approval Rating

$$\hat{p} = 0.52 \qquad n = 900 \qquad 95\%\, confidence >>> \alpha = 0.05$$

$$\hat{p} \pm z_{0.05/2} \cdot \sqrt{\frac{(\hat{p}(1-\hat{p})}{n}}$$

95% CI: $0.52 \pm 1.96\sqrt{\frac{0.52(1-0.52)}{900}}$

0.52 ± 0.033 ← Margin of error is 3.3%.

$0.49\, to\, 0.55$

Work problem #7 a, c, d.

Summary Page for Chapter 9

CHAPTER 9 Confidence Intervals

Confidence Intervals

- A $(1 - \alpha) \cdot 100\%$ confidence interval about μ with σ known is $\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ provided the population from which the sample was drawn is normal or the sample size is large $(n \geq 30)$.
- A $(1 - \alpha) \cdot 100\%$ confidence interval about μ with σ unknown is $\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$ provided the population from which the sample was drawn is normal or the sample size is large $(n \geq 30)$. *Note*: $t_{\alpha/2}$ is computed using $n - 1$ degrees of freedom.
- A $(1 - \alpha) \cdot 100\%$ confidence interval about p is $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ provided $n\hat{p}(1 - \hat{p}) \geq 10$.
- A $(1 - \alpha) \cdot 100\%$ confidence interval about σ^2 is $\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$ provided the population from which the sample was drawn is normal.

Sample Size

- To estimate the population mean with a margin of error E at a $(1 - \alpha) \cdot 100\%$ level of confidence requires a sample of size $n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$ rounded up to the next integer.
- To estimate the population proportion with a margin of error E at a $(1 - \alpha) \cdot 100\%$ level of confidence requires a sample of size $n = \hat{p}(1 - \hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$ rounded up to the next integer, where $\hat{p}$ is a prior estimate of the population proportion.
- To estimate the population proportion with a margin of error E at a $(1 - \alpha) \cdot 100\%$ level of confidence requires a sample of size $n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2$ rounded up to the next integer when no prior estimate of p is available.