




Chapter 9

Inferences Involving One Population



Chapter 9 Overview

- 9-1 Inferences about the Mean μ
(σ Unknown)
- 9-2 Inferences about the Binomial
Probability of Success



9.1 Inferences about the Mean μ (σ 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

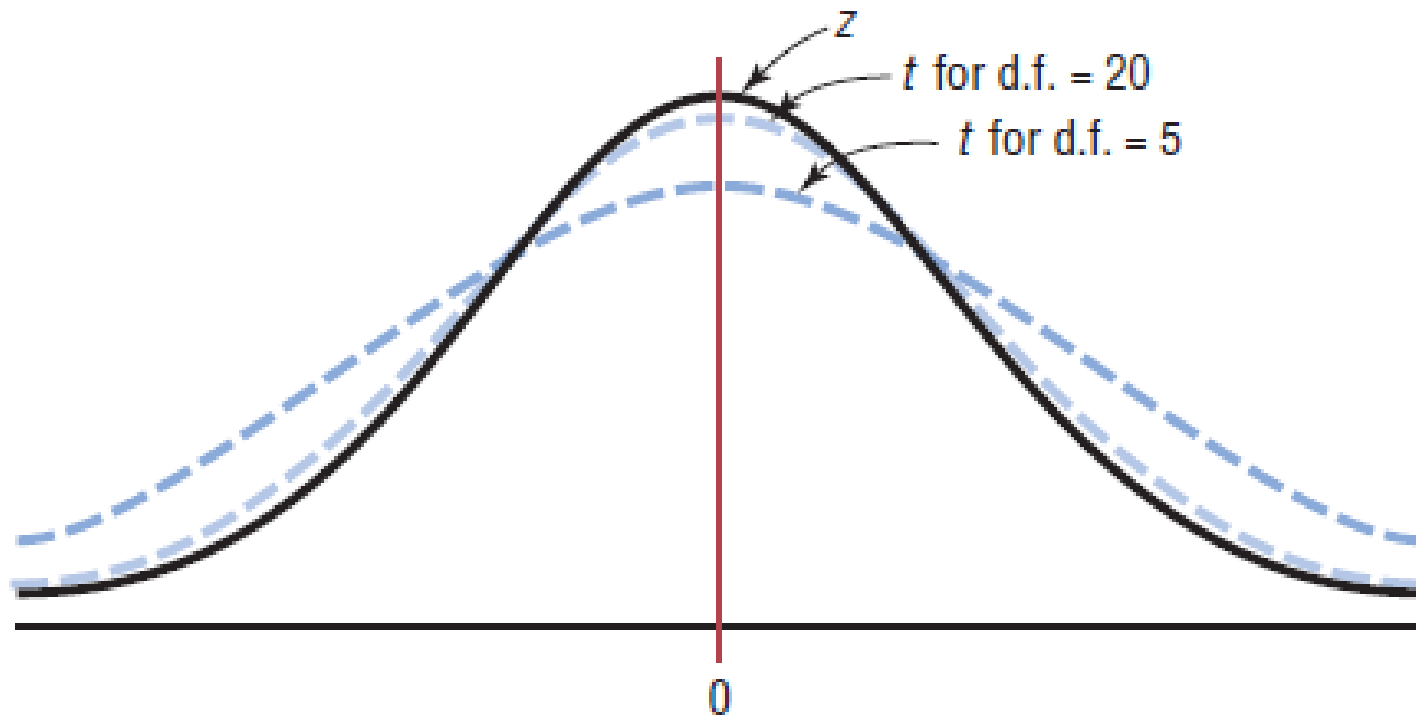
- The symbol d.f. will be used for **degrees of freedom**.
- The degrees of freedom for a confidence interval for the mean are found by subtracting 1 from the sample size. That is, $d.f. = n - 1$.
- Note: For some statistical tests used later in this book, the degrees of freedom are not equal to $n -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



Finding d.f. (degrees of freedom)

- Find d.f. (degrees of freedom) for the following:
 - ☐ A sample of size 25
 - ☐ A sample of size 53
 - ☐ A sample of size 3

Family of t curves



Ex.) 1 – Using the t distribution Table

Find the $t_{\alpha/2}$ value for a 90% confidence interval when the sample size is 10. Use the t distribution table in the formula sheet given.

Degrees of freedom are $d.f. = n - 1 = 9$.

d.f.	Confidence intervals	80%	90%	95%
	One tail, α	0.10	0.05	0.025
	Two tails, α	0.20	0.10	0.05
1		3.078	6.314	12.706
2		1.886	2.920	4.303
3		1.638	2.353	3.182
4		1.533	2.132	2.776
5		1.476	2.015	2.571
6		1.440	1.943	2.447
7		1.415	1.895	2.365
8		1.397	1.860	2.306
9		1.383	1.833	2.262
10		1.372	1.812	2.228
11		1.363	1.796	2.201



Finding t values

- Look up the following on the t table:
 - 90% and $df = 24$
 - 80% and $df = 3$
 - 0.025 in one tail and $df = 15$
 - 0.10 in two tails and $df = 18$



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 $n - 1$.

Ex.) 2 – Finding Confidence Interval for a Mean when σ is Unknown

Hemoglobin The average hemoglobin reading for a sample of 20 teachers was 16 grams per 100 milliliters, with a sample standard deviation of 2 grams. Find the 99% confidence interval of the true mean.

Since σ is unknown and s must replace it, the t distribution (Table 6) must be used for the confidence interval. Hence, with 19 degrees of freedom, $t_{\alpha/2} = 2.861$.

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

$$16 - 2.861 \left(\frac{2}{\sqrt{20}} \right) < \mu < 16 + 2.861 \left(\frac{2}{\sqrt{20}} \right)$$



Ex.) 2 – Finding Confidence Interval for a Mean when σ is Unknown

$$16 - 2.861 \left(\frac{2}{\sqrt{20}} \right) < \mu < 16 + 2.861 \left(\frac{2}{\sqrt{20}} \right)$$

$$16 - 1.28 < \mu < 16 + 1.28$$

$$14.62 < \mu < 17.28$$


One can be 99% confident that the population mean is between 14.62 and 17.28 grams per milliliter .



Ex.) 3 – Finding Confidence Interval for a Mean when σ is Unknown

Digital Camera Prices The prices (in dollars) for a particular model of digital camera with 6.0 megapixels and an optical 3X zoom lens are shown below for 10 online retailers. Estimate the true mean price for this particular model with 95% confidence.

225 240 215 206 211 210 193 250 225 202



Class Activity: Confidence Interval for the Population Mean when σ is unknown

Freshmen's GPA First-semester GPAs for a random selection of freshmen at a large university are shown below. Estimate the true mean GPA of the freshman class with 99% confidence. .

1.9	3.2	2.0	2.9	2.7	3.3
2.8	3.0	3.8	2.7	2.0	1.9
2.5	2.7	2.8	3.2	3.0	3.8
3.1	2.7	3.5	3.8	3.9	2.7
2.0	2.8	1.9	4.0	2.2	2.8
2.1	2.4	3.0	3.4	2.9	2.1

How does this differ from an estimate where we assume that we know the population standard deviation and it is $\sigma = 0.62$?



t Test for a Mean

The *t* test is a statistical test for the mean of a population and is used when the population is normally or approximately normally distributed, σ is unknown.

The formula for the *t* test is

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

The degrees of freedom are d.f. = $n - 1$.

Ex.) 4 – Finding Critical t value

Find the critical t value for $\alpha = 0.05$ with d.f. = 16 for a right-tailed t test.

Find the 0.05 column in the top row and 16 in the left-hand column.
The critical t value is +1.746.

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							
4							
5							
⋮							
14							
15							
16				1.746			



Ex.) 5 – Finding Critical t value

Find the critical t value for $\alpha = 0.01$ with d.f. = 22 for a left-tailed test.

Ex.) 6 – Finding Critical t value

Find the critical value for $\alpha = 0.10$ with d.f. = 18 for a two-tailed t test.



Ex.) 7 – Using t Test for a Mean

Veterinary Expenses of Cat Owners According to the American Pet Products Manufacturers Association, cat owners spend an average of \$179 annually in routine veterinary visits. A random sample of local cat owners revealed that 10 randomly selected owners spent an average of \$205 with $s = \$26$. Is there a significant statistical difference at $\alpha = 0.01$?

Step 1: State the hypotheses and identify the claim.

$$H_0: \mu = 179 \text{ and } H_1: \mu \neq 179 \text{ (claim)}$$

Step 2: Find the critical value.

The critical values are ± 3.250 for $\alpha = 0.01$ and d.f. = 9.



Ex.) 7 – Using t Test for a Mean

Veterinary Expenses of Cat Owners According to the American Pet Products Manufacturers Association, cat owners spend an average of \$179 annually in routine veterinary visits. A random sample of local cat owners revealed that 10 randomly selected owners spent an average of \$205 with $s = \$26$. Is there a significant statistical difference at $\alpha = 0.01$?

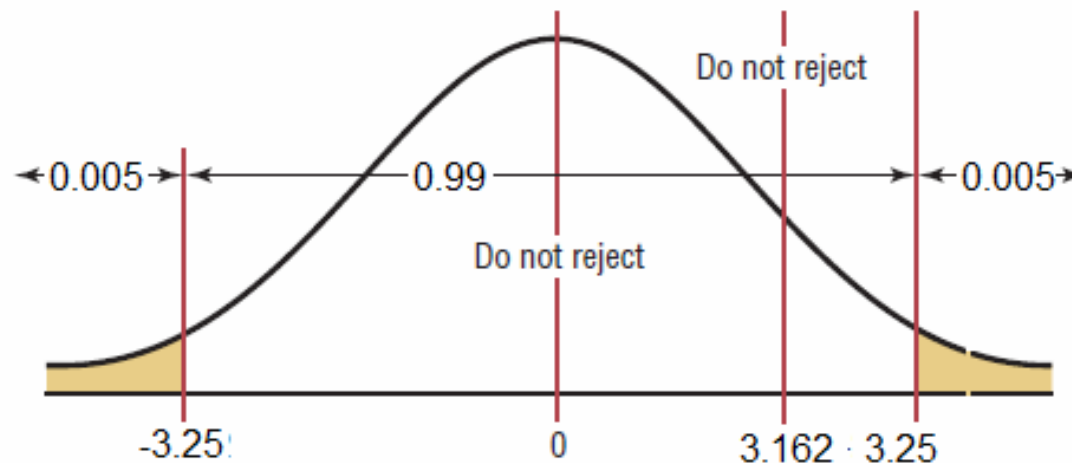
Step 3: Find the test value.

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{205 - 179}{26/\sqrt{10}} = 3.162$$

Ex.) 7 – Using t Test for a Mean

Step 4: Make the decision.

Do not reject the null hypothesis since 3.162 falls in the non-critical region.



Step 5: Summarize the results.

There is not enough evidence to reject the claim that the average expense is \$179.

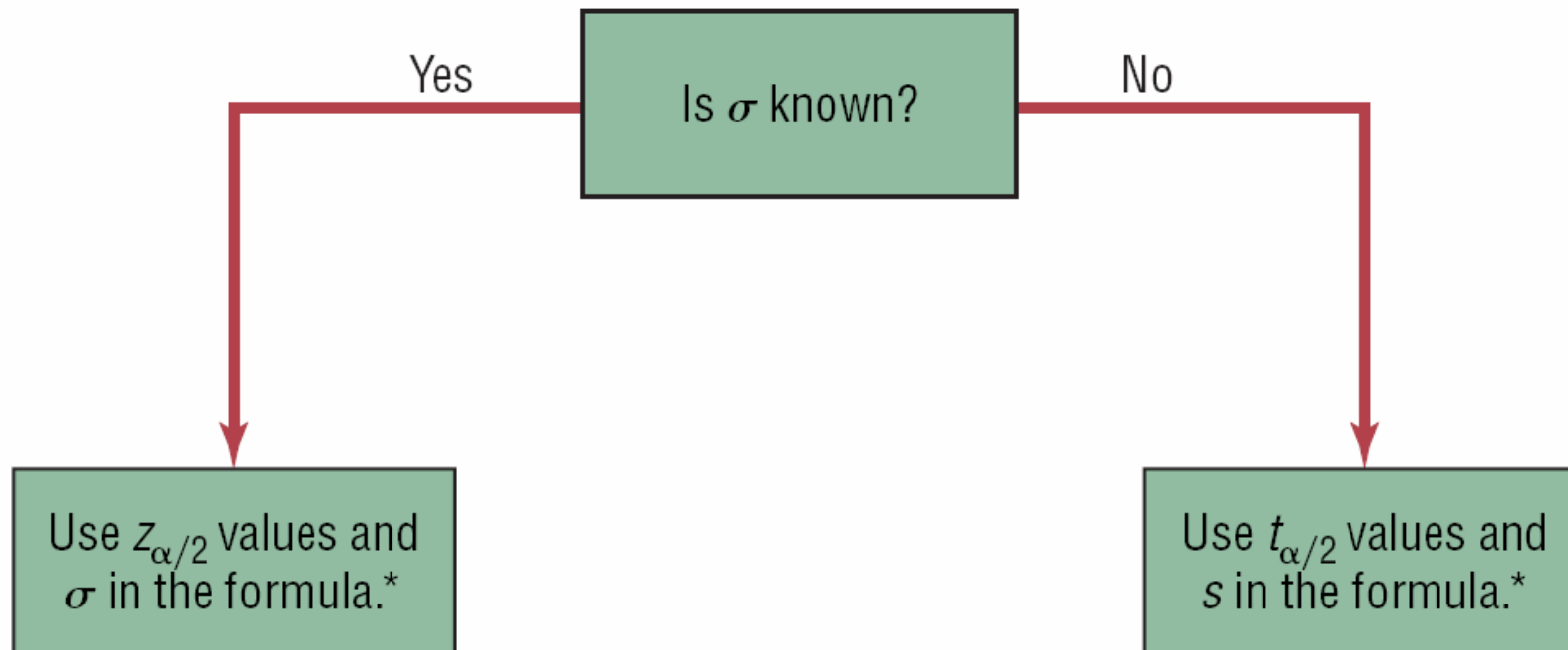


Ex.) 8 – Using t Test for a Mean

Heights of Tall Buildings A researcher estimates that the average height of the buildings of 30 or more stories in a large city is at least 700 feet. A random sample of 10 buildings is selected, and the heights in feet are shown. At $\alpha = 0.025$, is there enough evidence to reject the claim?

485	511	841	725	615
520	535	635	616	582

Whether to use z or t



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

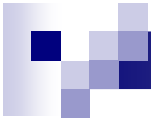


Class Activity: t Test for a Mean

Doctor Visits A report by the Gallup Poll stated that on average a woman visits her physician 5.8 times a year. A researcher randomly selects 20 women and obtained these data.

3	2	1	3	7	2	9	4	6	6
8	0	5	6	4	2	1	3	4	1

At $\alpha = 0.05$ can it be concluded that the average is still 5.8 visits per year?



9.2 Inferences about the Binomial Probability of Success

p = population proportion

p' = sample proportion

For a sample proportion,

$$p' = \frac{X}{n}, \quad q' = 1 - p'$$

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



Ex.) 9 – Finding Proportions

A recent poll found that 329 out of 763 adults would travel to space in their lifetime given the chance. Find p' and q' , where p' is the proportion of people who would travel to space given the chance.

Since $X = 329$ and $n = 763$,

$$p' = \frac{X}{n} = \frac{329}{763} = 0.431 = 43.1\%$$

$$q' = 1 - p' = 1 - 0.431 = 0.569 = 56.9\%$$



Formula for a Specific Confidence Interval for a Proportion

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

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

Rounding Rule: Round off to three decimal places.

Ex.) 10 – Finding Confidence Interval for a Proportion

A recent poll found that 329 out of 763 adults would travel to space in their lifetime given the chance. Find the 92% confidence interval of the true proportion of adults who would like to travel to space.

$$p' = 0.431, \quad q' = 0.569, \quad z = 1.75$$

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

$$0.431 - 1.75 \sqrt{\frac{0.431 \cdot 0.569}{763}} < p < 0.431 + 1.75 \sqrt{\frac{0.431 \cdot 0.569}{763}}$$

$$0.431 - 0.031 < p < 0.431 + 0.031$$

$$0.400 < p < 0.462$$

We are 92% confident that the percentage of adults who would like to travel to space is between 40.0% and 46.2%.



Determining the Sample Size

- By using the maximum error part of the confidence interval formula, it is possible to determine the **size of the sample** that must be taken in order to estimate p with a desired accuracy.
- Here is the formula for the **maximum error of estimate for a proportion**:

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



Determining the Sample Size

- To determine the sample size from this formula, we must decide on the quality we want for our final confidence interval.
- This quality is measured in two ways: the level of confidence and the preciseness (narrowness) of the interval.
- The level of confidence we establish will in turn determine the confidence coefficient, $z(\alpha/2)$.



Determining the Sample Size

- The desired preciseness will determine the maximum error of estimate, E . (Remember that we are estimating p , the binomial probability; therefore, E will typically be expressed in hundredths.)

Sample Size for $1 - \alpha$ Confidence Interval of p

$$n = \frac{[z(\alpha/2)]^2 \cdot p^* \cdot q^*}{E^2} \quad (9.8)$$

- where p^* and q^* are provisional values of p and q used for planning



Determining the Sample Size

- By inspecting formula (9.8), we can observe that three components determine the sample size:

1. The level of confidence $[1 - \alpha]$, which determines the confidence coefficient, $z(\alpha/2)$

- 2. The provisional value of p (p^* determines the value of q^*)

- 3. The maximum error, E

- An increase or decrease in one of these three components affects the sample size.



Determining the Sample Size

- If the level of confidence is increased or decreased (while the other components are held constant), then the sample size will increase or decrease, respectively.
- If the product of p^* and q^* is increased or decreased (with other components held constant), then the sample size will increase or decrease, respectively. (The product $p^* \cdot q^*$ is largest when $p^* = 0.5$ and decreases as the value of p^* becomes further from 0.5.)
- An increase or decrease in the desired maximum error will have the opposite effect on the sample size, since E appears in the denominator of the formula.



Determining the Sample Size

- If no provisional values for p and q are available, then use $p^* = 0.5$ and $q^* = 0.5$. Using $p^* = 0.5$ is safe because it gives the largest sample size of any possible value of p .
- Using $p^* = 0.5$ works reasonably well when the true value is “near 0.5” (say, between 0.3 and 0.7); however, as p gets nearer to either 0 or 1, a sizable overestimate in sample size will occur.



Ex.) 11 – Finding the Minimum Sample Size Needed to Estimate a Population Proportion

Vitamins for Women A medical researcher wishes to determine the percentage of females who take vitamins. He wishes to be 99% confident that the estimate is within 2 percentage points of the true proportion. A recent study of 180 females showed that 25% took vitamins.

a. How large should the sample size be?

$$n = \frac{[z_{\alpha/2}]^2 p^* q^*}{E^2} = \frac{2.58^2 * .25 * .75}{.02^2} = 3120.1875 \approx 3121$$

The researcher should use a sample of at least 3121 women.



Ex.) 11 – Finding the Minimum Sample Size Needed to Estimate a Population Proportion

- b.* If no estimate of the sample proportion is available, how large should the sample be?

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

$$n = \frac{[z_{\alpha/2}]^2 p^* q^*}{E^2} = \frac{2.58^2 * .5 * .5}{.02^2} = 4160.25 \approx 4161$$

The researcher should use a sample of at least 4161 women.



Class Activity: Finding Confidence Interval for a Proportion

Freshmen's GPA First-semester GPAs for a random selection of freshmen at a large university are shown below. Estimate the true mean GPA of the freshman class with 99% confidence. .

1.9	3.2	2.0	2.9	2.7	3.3
2.8	3.0	3.8	2.7	2.0	1.9
2.5	2.7	2.8	3.2	3.0	3.8
3.1	2.7	3.5	3.8	3.9	2.7
2.0	2.8	1.9	4.0	2.2	2.8
2.1	2.4	3.0	3.4	2.9	2.1

From this previous class activity, find a 95% confidence interval for the proportion of GPAs greater than 3.0.



z Test for a Proportion

Since a normal distribution can be used to approximate the binomial distribution when $np \geq 5$ and $nq \geq 5$, the standard normal distribution can be used to test hypotheses for proportions.

The formula for the z test for a proportion is

$$z = \frac{p' - p}{\sqrt{pq / n}}$$

where

$$p' = \frac{\bar{X}}{n} \quad (\text{sample proportion})$$

p = population proportion

n = sample size



Procedure for Hypothesis Testing of Binomial Data

Step 1 The Set-Up:

- **a. Describe the population parameter of interest.**
- **b. State the null hypothesis (H_o) and the alternative hypothesis (H_a).**

Step 2 The Hypothesis Test Criteria:

- **a. Check the assumptions.**
 - sample is independent & randomly selected
 - can the normal approximation to the binomial distribution be used (ie $np > 5$? $nq > 5$?)
- **b. Identify the probability distribution and the test statistic to be used.**



Procedure for Hypothesis Testing of Binomial Data cont'd

- c. Determine the level of significance

Step 3 The Sample Evidence:

- a. Collect the sample information
- b. Calculate the value of the test statistic.

Step 4 The Probability Distribution:

- a. Determine the critical region and critical value (s).
- b. Determine whether or not the calculated test statistic is in the critical region.

Step 5 The Results:

- a. State the decision about H_0
- b. State the conclusion about H_a



Ex.) 12 – z Test for a Proportion

Answering Machine Ownership It has been reported in an almanac that 78% of Americans own an answering machine. A random sample of 143 college professors at small liberal arts schools revealed that 100 owned an answering machine. At $\alpha = 0.05$, test the claim that the percentage is the same as stated in the almanac.

Step 1: State the hypotheses and identify the claim.

$$H_0: p = 0.78 \text{ (claim) and } H_1: p \neq 0.78$$

Step 2: Find the critical value.

Since $\alpha = 0.05$ and the test is a two-tailed test, the critical value is $z = \pm 1.96$.



Ex.) 12 – z Test for a Proportion

Answering Machine Ownership It has been reported in an almanac that 78% of Americans own an answering machine. A random sample of 143 college professors at small liberal arts schools revealed that 100 owned an answering machine. At $\alpha = 0.05$, test the claim that the percentage is the same as stated in the almanac.

Step 3: Compute the test value.

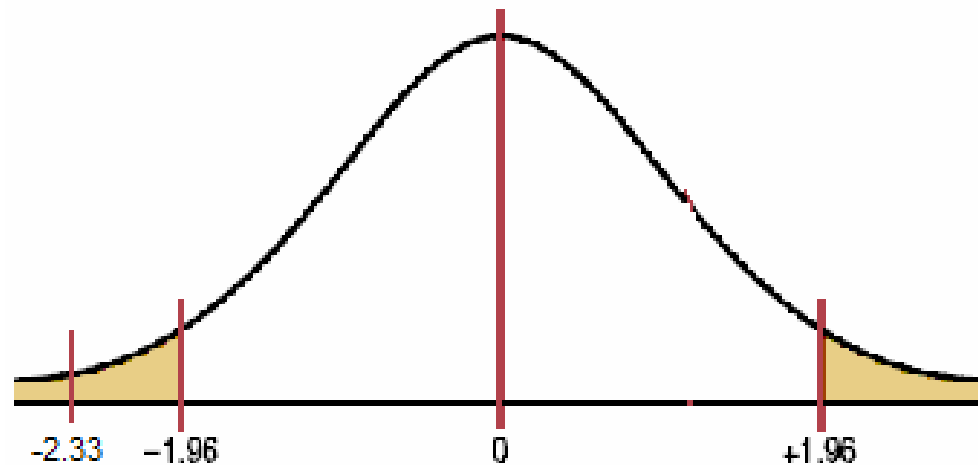
$$p' = \frac{\bar{X}}{n} = \frac{100}{143} = 0.6993$$

$$z = \frac{p' - p}{\sqrt{pq/n}} = \frac{0.6993 - 0.78}{\sqrt{(0.78)(0.22)/143}} = -2.33$$

Ex.) 12 – z Test for a Proportion

Step 4: Make the decision.

Reject the null hypothesis since the test value falls in the critical region.



Step 5: Summarize the results.

There is enough evidence to reject the claim that the percentage is 78%.



Class Activity: z Test for a Proportion

First-Class Airline Passengers In the past, 20% of all airline passengers flew first class. In a sample of 15 passengers, 5 flew first class. At $\alpha = 0.10$, can you conclude that the proportions have changed?