LAYING the FOUNDATION®

AP* Statistics Review

Confidence Intervals

Teacher Packet

A **Confidence Interval** is an interval that is computed from sample data and provides a range of plausible values for a population parameter.
A **Confidence Level** is a number that provides information on how much "confidence" we have in the method used to construct a confidence interval estimate. This level specifies the percentage of all possible samples that produce an interval containing the true value of the population parameter.

Constructing a Confidence Interval
The steps listed below should be followed when asked to calculate a confidence interval.

1. Identify the population of interest and define the parameter of interest being estimated.
2. Identify the appropriate confidence interval by name or formula.
3. Verify any conditions (assumptions) that need to be met for that confidence interval.
4. Calculate the confidence interval.
5. Interpret the interval in the context of the situation.

The general formula for a confidence interval calculation is:

Statistic ± (critical value) · (standard deviation of statistic)

The critical value is determined by the confidence level.
The type of statistic is determined by the problem situation. Here we will discuss two types of statistics: mean and proportion.

Confidence Interval for Proportions: (1-proportion z-interval)
We are finding an interval that describes the population proportion ($p \text{ or } \pi$).
The general formula uses the sample proportion ($\hat{p}$) and the sample size (n).

Statistic ± (critical value) · (standard deviation of statistic)

$$\hat{p} \pm \left(z^*\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

The conditions that need to be met for this procedure are:
1. The sample is a simple random sample.
2. The population is large relative to the sample.
   $10n < N$ (N = the size of the population)
3. The sampling distribution of the sample proportion is approximately normal.
   $np \geq 10$
   $n(1-p) \geq 10$

Complete the calculations after showing that the conditions are met.
Write the answer in the context of the original problem.

Sample question:

The owner of a popular chain of restaurants wishes to know if completed dishes are being delivered to the customer's table within one minute of being completed by the chef. A random sample of 75 completed dishes found that 60 were delivered within one minute of completion. Find the 95% confidence interval for the true population proportion.

1. Identify the population of interest and define the parameter of interest being estimated.

The population of interest is the dishes that are being served at this chain of restaurants.

p = the population proportion of dishes that are served within one minute of completion.

$\hat{p} = \frac{60}{75} = 0.8$ = the sample proportion of dishes that are served within one minute of completion.

$n = 75$ is the sample size.

2. Identify the appropriate confidence interval by name or formula.

We will use a 95% confidence z-interval for proportions.

3. Verify any conditions (assumptions) that need to be met for that confidence interval.

The problem states that this is a simple random sample.

$10n < N$

$10(75) < N$

$750 < N$

It is reasonable to assume that a popular chain of restaurants will serve more than 750 dishes.

$n\hat{p} \geq 10$	$n(1-\hat{p}) \geq 10$
$75(0.8) = 60$	$75(1-.8) = 15$
$60 \geq 10$	$15 \geq 10$

It is reasonable to use a normal model.

4. Calculate the confidence interval.

At a 95% CI, the critical value is $z^* = 1.96$.

This value is found on the last row of the t-distribution table.

$$0.8 \pm 1.96\sqrt{\frac{0.8(1-0.8)}{75}}$$

$$0.8 \pm 0.091$$

$$(0.709, 0.891)$$

5. Interpret the interval in the context of the situation.

We are 95% confident that the true proportion of dishes that are served within one minute of completion for this chain of restaurants is between 0.709 and 0.891.

<u>Confidence Interval for Means with σ known:</u> (z-interval)
We are finding an interval that describes the population mean (μ).
The general formula uses the sample mean ($\bar{x}$), the population standard deviation (σ) and the sample size (n).

Statistic ± (critical value) · (standard deviation of statistic)

$$\bar{x} \pm (z^*)\left(\frac{\sigma}{\sqrt{n}}\right)$$

The conditions that need to be met for this procedure are:
1. The sample is a simple random sample.
2. The population is normal or $n \geq 30$
3. The population standard deviation (σ) is known.

Complete the calculations after showing that the conditions are met.
Write your answer in the context of the original problem.

<u>Sample question:</u>

An asbestos removal company places great importance on the safety of their employees. The protective suits that the employees wear are designed to keep asbestos particles off the employee's body. The owner is interested in knowing the average amount of asbestos particles left on the employee's skin after a days work. A random sample of 100 employees had skin tests after removing their protective suit. The average number of particles found was .481 particles per square centimeter. Assuming that the population standard deviation is 0.35 particles per square centimeter, calculate a 95% confidence interval for the number of particles left on the employee's skin.

<u>Solution:</u>

The population of interest is the employees of the asbestos removal company.
μ = the population mean of particles per square centimeter after a days work wearing the protective suit.
$\bar{x} = 0.481$ = the sample mean of the number of particles per square centimeter.
$n = 100$ is the sample size.
$\sigma = 0.35$ =population standard deviation.

We will use a 95% confidence z-interval for means (z-interval).
The problem states that this is a simple random sample.
n = 100 is greater than 30 so we can assume that use of a normal model is reasonable.

At a 95% CI, the critical value is $z^* = 1.96$.
This value is found on the last row of the t-distribution table.

$$0.481 \pm 1.96\left(\frac{0.35}{\sqrt{100}}\right)$$

$$0.481 \pm 0.069$$

$$(0.412, 0.550)$$

We are 95% confident that the true mean number of particles of asbestos found on the skin of an employee after a days work is between 0.412 and 0.550 particles per square centimeter.

<u>Confidence Interval for Means with σ unknown:</u> (t-interval)

You will be finding an interval that will describe the population mean (μ).

The general formula will use the sample mean ($\bar{x}$), the sample standard deviation (s), the sample size (n), and the degrees of freedom ($n-1$).

Statistic ± (critical value) · (standard deviation of statistic)

$$\bar{x} \pm (t^*)\left(\frac{s}{\sqrt{n}}\right)$$

The conditions that need to be met for this procedure are:

1. The sample is a simple random sample.
2. The population is approximately normal (graphical support required) or $n \geq 40$
3. The population standard deviation (σ) is unknown.

Complete the calculations after showing that the conditions are met.
Write your answer in the context of the original problem.

This situation (where the population standard deviation is unknown) is much more realistic that the previous case.

<u>Sample question:</u>

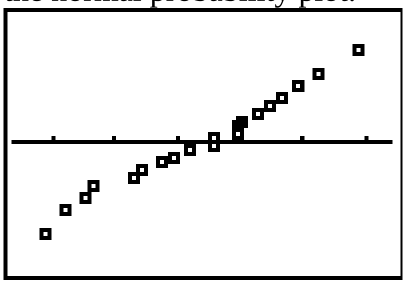
A biology student at a major university is writing a report about bird watchers. She has developed a test that will score the abilities of a bird watcher to identify common birds. She collects data from a random sample of people that classify themselves as bird watchers (data shown below). Find a 90% confidence interval for the mean score of the population of bird watchers.

4.5	*9.1*	*8*	*5.9*	*7.0*	*5.2*	*7.3*	*7.0*	*6.6*	*5.1*
7.6	*8.2*	*6.4*	*4.8*	*5.8*	*6.2*	*8.5*	*7.3*	*7.8*	*7.4*

Solution:

The population of interest is people that classify themselves as bird watchers.
μ = the population mean score on the bird identification ability test.
$\bar{x} = 6.785$ = the sample mean of the scores on the ability test.
$s = 1.2828$ = sample standard deviation of test scores.
$n = 20$ is the sample size.
df = 20 – 1 = 19 degrees of freedom.

We will use a 90% confidence t-interval for means (t-interval).
The problem states that this is a simple random sample.
The sample size, 20, is smaller than 40 so we will assess normality by looking at the normal probability plot.



The normal probability plot appears to be linear so we will assume an approximately normal distribution of scores.
The population standard deviation is unknown.

At a 90% CI, the critical value is $t^* = 1.729$.
This value is found on the t-distribution table using 19 degrees of freedom.

$$6.785 \pm 1.729\left(\frac{1.2828}{\sqrt{20}}\right)$$

$$6.785 \pm 0.49595$$

$$(6.289, 7.281)$$

We are 90% confident that the true mean score on the bird identification ability test of the population of persons that classify themselves as bird watchers is between 6.289 and 7.281.

Summary of Confidence Intervals with One Sample

Confidence Interval Type	**Formula**	**Conditions**	**Calculator Test**
Proportions	$\hat{p} \pm (z^*)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	1. The sample is a simple random sample. 2. The population is large relative to the sample $10n < N$ 3. $np \geq 10$ $n(1-p) \geq 10$	1-PropZInterval
Means (σ known)	$\bar{x} \pm (z^*)\left(\frac{\sigma}{\sqrt{n}}\right)$	1. The sample is a simple random sample. 2. The population is normal or $n \geq 30$ 3. The population standard deviation (σ) is known.	ZInterval
Means (σ unknown)	$\bar{x} \pm (t^*)\left(\frac{s}{\sqrt{n}}\right)$	1. The sample is a simple random sample. 2. The population is approximately normal (graphical support required) or $n \geq 40$ 3. The population standard deviation (σ) is unknown.	TInterval

Confidence Intervals with TWO Samples

We may be asked to work a confidence interval problem to estimate the difference of two population parameters. The process is very similar to that of a one sample confidence interval. We must make sure that the conditions are met for both samples and that the samples are independent of each other.

The formulas for the two sample confidence intervals are as follows:

Confidence Interval Type	Formula	Conditions	Calculator Test
Difference of Proportions $(p_1 - p_2)$	$\hat{p}_1 - \hat{p}_2 \pm (z^*)\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	1. The samples are independently selected simple random samples. 2. Each population is large relative to the sample $10n_1 < N_1$ $10n_2 < N_2$ 3. $n_1 p_1 \geq 10$ $n_1(1-p_1) \geq 10$ $n_2 p_2 \geq 10$ $n_2(1-p_2) \geq 10$	2-PropZInterval
Difference of Means (σ known) $(\mu_1 - \mu_2)$	$\bar{x}_1 - \bar{x}_2 \pm (z^*)\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	1. The samples are independently selected simple random samples. 2. Both populations are normal or $n_1 \geq 30$ $n_2 \geq 30$ 3. Both population standard deviations (σ_1 *and* σ_2) are known.	2-samp-ZInterval
Difference of Means (σ unknown) $(\mu_1 - \mu_2)$	$\bar{x}_1 - \bar{x}_2 \pm (t^*)\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ Note: The formula for degrees of freedom is $df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$ This formula is not provided on the formula chart. A more conservative quantity used by some statisticians is the smaller of the two individual degrees of freedom; either $n_1 - 1$ or $n_2 - 1$.	1. The samples are independently selected simple random samples. 2. Both populations are approximately normal (graphical support is required) or $n_1 \geq 40$ $n_2 \geq 40$ 3. Both population standard deviations (σ_1 *and* σ_2) are unknown.	2-samp-TInterval

Sample question:

Two popular strategy video games, AE and C, are known for their long play times. A popular game review website is interested in finding the mean difference in play time between these games. The website selects a random sample of 43 gamers to play AE and finds their sample mean play time to be 3.6 hours with a standard deviation of 0.9 hours. The website also selected a random sample of gamers to test the game C. There test included 40 gamers with a sample mean of 3.1 hours and a standard deviation of 0.4 hours. Find the 98% confidence interval for the difference $\mu_{AE} - \mu_C$.

Solution:

The population of interest is play time of games AE and C.

μ_{AE} = the population mean play time for game AE.

μ_C = the population mean play time for game C.

$\bar{x}_{AE} = 3.6$ = the sample mean of play time for game AE.

$\bar{x}_C = 3.1$ = the sample mean of play time for game C.

$s_{AE} = 0.9$ = sample standard deviation play time for game AE.

$s_C = 0.4$ = sample standard deviation play time for game C.

$n_{AE} = 43$ is the sample size of the players of game AE.

$n_C = 40$ is the sample size of the players of game C.

df = 43 – 1 = 42 degrees of freedom (using the simplified approximation for df).

We will use a 98% confidence t-interval for the difference of means (2-sample-t-interval).
The problem states that each is a simple random sample.
The sample sizes are 43 and 40 which are both at least 40.
The population standard deviation is unknown.

At a 98% CI, the critical value is $t^* = 2.457$.
This value is found on the t-distribution table using 40 degrees of freedom (because the table does not have a value for 42).

$$(3.6 - 3.1) \pm 2.423\sqrt{\frac{0.9^2}{43} + \frac{0.4^2}{40}}$$

$$0.5 \pm 0.3662$$

$$(0.1338, 0.8662)$$

We are 98% confident that the true difference in mean play time between games AE and C falls between 0.1338 and 0.8662 hours.

Note: The calculator gives a result of (0.1386, 0.8614) for this confidence interval. The difference in these answers is due to its use of 58.872 degrees of freedom. Our answer has a slightly wider spread, and thus is more conservative.

Other Confidence Interval Topics:

Statistic ± (critical value) · (standard deviation of statistic)

Margin of Error (ME) – the margin of error is the value of the critical value times the standard deviation of the statistic. It is the plus or minus part of the confidence interval.

Some problems might ask you to determine the sample size required given a margin of error. This requires a little algebra to work backward through the equations. The equations are listed below.

Confidence Interval Type	Formula for finding the sample size within a margin of error ME
1-sample proportion	$n = \hat{p}(1-\hat{p})\left(\frac{z^*}{ME}\right)^2$ Note: if $\hat{p}$ is not given in this type of problem, a conservative value to use if 0.5.
1-sample mean (z-interval)	$n = \left(\frac{z^*\sigma}{ME}\right)^2$

Keep in mind the effects of changing the confidence level. A large confidence level (say 99% as compared to 90%) produces a larger margin of error. To be more confident we must include more values in our range.

Do not confuse the meaning of a confidence level: A 95% confidence level means that if we repeated the sampling process many times, the resulting confidence interval would capture the true population parameter 95% of the time.

Multiple Choice Questions on Confidence Intervals

1. A random sample of 100 visitors to a popular theme park spent an average of $142 on the trip with a standard deviation of $47.5. Which of the following would the 98% confidence interval for the mean money spent by all visitors to this theme park?

   (A) ($130.77, $153.23)
   (B) ($132.57, $151.43)
   (C) ($132.69, $151.31)
   (D) ($140.88, $143.12)
   (E) ($95.45, $188.55)

2. How large of a random sample is required to insure that the margin of error is 0.08 when estimating the proportion of college professors that read science fiction novels with 95% confidence?

   (A) 600
   (B) 300
   (C) 150
   (D) 75
   (E) 25

3. A quality control specialist at a plate glass factory must estimate the mean clarity rating of a new batch of glass sheets being produced using a sample of 18 sheets of glass. The actual distribution of this batch is unknown, but preliminary investigations show that a normal approximation is reasonable. The specialist decides to use a t-distribution rather than a z-distribution because

   (A) The z-distribution is not appropriate because the sample size is too small.
   (B) The sample size is large compared to the population size.
   (C) The data comes from only one batch.
   (D) The variability of the batch is unknown.
   (E) The t-distribution results in a narrower confidence interval.

4. An independent random sample of 200 college football players and 150 college basketball players in a certain state showed that 65% of football players received academic tutors while 58% of basketball players received academic tutors. Which of the following is a 90 percent confidence interval for the difference in the proportion of football players that received tutors and the proportion of basketball players that received tutors for the population of this state?

(A) $(.65-.58)\pm 1.96\sqrt{(0.65)(0.58)\left(\frac{1}{200}+\frac{1}{150}\right)}$

(B) $(.65-.58)\pm 1.645\sqrt{(0.65)(0.58)\left(\frac{1}{200}+\frac{1}{150}\right)}$

(C) $(.65-.58)\pm 1.96\sqrt{\frac{(0.65)(0.35)}{200}+\frac{(0.58)(0.42)}{150}}$

(D) $(.65-.58)\pm 1.645\sqrt{\frac{(0.65)(0.35)}{200}+\frac{(0.58)(0.42)}{150}}$

(E) $(.65-.58)\pm 1.645\sqrt{(0.435)(0.867)\left(\frac{1}{200}+\frac{1}{150}\right)}$

5. The board of directors at a city zoo is considering using commercial fast food restaurants in their zoo rather than the current eateries. They are concerned that major donors to the zoo will not approve of the proposed change. Of the 280 major donors to the zoo, a random sample of 90 is asked "Do you support the zoo's decision to use commercial fast food restaurants in the zoo?" 50 of the donors said no, 38 said yes, and 2 had no opinion on the matter. A large sample z-interval, $\hat{p}\pm(z^*)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, was constructed from these data to estimate the proportion of the major donors who support using commercial fast food restaurants in the zoo. Which of the following statements is correct for this confidence interval?

(A) This confidence interval is valid because a sample size of more than 30 was used.
(B) This confidence interval is valid because no conditions are required for constructing a large sample confidence interval for proportions.
(C) This confidence interval is not valid because the sample size is too large compared to the population size.
(D) This confidence interval is not valid because the quantity $n\hat{p}$ is too small.
(E) This confidence interval is not valid because "no opinion" was allowed as a response.

6. A research and development engineer is preparing a report for the board of directors on the battery life of a new cell phone they have produced. At a 95% confidence level, he has found that the battery life is 3.2 ± 1.0 days. He wants to adjust his findings so the margin of error is as small as possible. Which of the following will produce the smallest margin of error?

(A) Increase the confidence level to 100%. This will assure that there is no margin of error.
(B) Increase the confidence level to 99%.
(C) Decrease the confidence level to 90%.
(D) Take a new sample from the population using the exact same sample size.
(E) Take a new sample from the population using a smaller sample size.

7. A biologist has taken a random sample of a specific type of fish from a large lake. A 95 percent confidence interval was calculated to be 6.8 ± 1.2 pounds. Which of the following is true?

(A) 95 percent of all the fish in the lake weigh between 5.6 and 8 pounds.
(B) In repeated sampling, 95 percent of the sample proportions will fall within 5.6 and 8 pounds.
(C) In repeated sampling, 95% of the time the true population mean of fish weights will be equal to 6.8 pounds.
(D) In repeated sampling, 95% of the time the true population mean of fish weight will be captured in the constructed interval.
(E) We are 95 percent confident that all the fish weigh less than 8 pounds in this lake.

8. A polling company is trying to estimate the percentage of adults that consider themselves happy. A confidence interval based on a sample size of 360 has a larger than desired margin of error. The company wants to conduct another poll and obtain another confidence interval of the same level but reduce the error to one-third the size of the original sample. How many adults should they now interview?

(A) 40
(B) 180
(C) 720
(D) 1080
(E) 3240

9. A researcher is interested in determining the mean energy consumption of a new compact florescent light bulb. She takes a random sample of 41 bulbs and determines that the mean consumption is 1.3 watts per hour with a standard deviation of 0.7. When constructing a 97% confidence interval, which would be the most appropriate value of the critical value?

(A) 1.936
(B) 2.072
(C) 2.250
(D) 2.704
(E) 2.807

10. A 98 percent confidence interval for the mean of a large population is found to be 978 ± 25. Which of the following is true?

(A) 98 percent of all observations in the population fall between 953 and 1003
(B) The probability of randomly selecting an observation between 953 and 1003 from the population is 0.98
(C) If the true population mean is 950, then this sample mean of 978 would be unlikely to occur.
(D) If the true population mean is 990, then this sample mean of 978 would be unexpected.
(E) If the true population mean is 1006, then this confidence interval must have been calculated incorrectly.

Free Response Questions on Confidence Intervals

Free Response 1.

A random sample of 9^{th} grade math students was asked if they prefer working their math problems using a pencil or a pen. Of the 250 students surveyed, 100 preferred pencil and 150 preferred pen.

(a) Using the results of this survey, construct a 95 percent confidence interval for the proportion of 9^{th} grade students that prefer to work their math problems in pen.

(b) A school newspaper reported on the results of this survey by saying, “Over half of ninth-grade math students prefer to use pen on their math assignments.” Is this statement supported by your confidence interval? Explain.

Free Response 2.

A major city in the United States has a large number of hotels. During peak travel times throughout the year, these hotels use a higher price for their rooms. A travel agent is interested in finding the difference of the average cost of a hotel rooms from the peak season to the off season. He takes a random sample of hotel room costs during each of these seasons. Plots of both samples of data indicate that the assumption of normality is not unreasonable.

Season	Cost	Standard Deviation	Sample Size
Peak Season	$245	$45	33
Off Season	$135	$65	38

(a) Construct a 95 percent confidence interval for the difference of the mean cost of hotel rooms from peak season to the off season.

(b) One particular hotel has an off season rate of $88 and a peak season rate of $218. Based on your confidence interval, comment on the price difference of this hotel.

Key to Confidence Interval Multiple Choice

1. A	Use the t-interval
2. C	Make a conservative assumption that the proportion is 0.5
3. D	The standard deviation of the population (batch) is not known.
4. D	2-prop-z-interval
5. C	Conditions of 1-proportion-z-interval
6. C	Decreasing the confidence level decreases the margin of error
7. D	Correct interpretation of the confidence level
8. E	Increasing the sample size by 9 decreases the error by one-third
9. C	Estimates from the t-table or invnorm(.985) produce this result.
10. C	Reasonable alternative interpretation of a confidence interval

Rubric for Confidence Interval Free Response

Free Response 1. **Solution**

Part (a)

1. Identify the population of interest and define the parameter of interest being estimated.

The population of interest is all ninth grade math students.

p = the population proportion of ninth grade math students that prefer to use pen on their math work.

$\hat{p} = \frac{150}{250} = 0.6$ = the sample proportion of ninth grade math students that prefer to use pen.

$n = 250$ is the sample size.

2. Identify the appropriate confidence interval by name or formula.

We will use a 95% confidence z-interval for proportions.

3. Verify any conditions (assumptions) that need to be met for that inference procedure.

The problem states that this is a simple random sample.

$10n < N$

$10(250) < N$

$2500 < N$

It is reasonable to assume that there are more than 2500 ninth grade math students.

$n\hat{p} \geq 10$	$n(1-\hat{p}) \geq 10$
$250(0.6) = 150$	$250(1-.6) = 100$
$150 \geq 10$	$100 \geq 10$

It is reasonable to use the normal approximation.

4. Calculate the confidence interval.

At a 95% CI, the critical value is $z^* = 1.96$.

$$0.6 \pm 1.96\sqrt{\frac{0.6(1-0.6)}{250}}$$

$$0.6 \pm 0.0607$$

$$(0.5393, 0.6607)$$

5. **Interpret your results in the context of the situation.**
   We are 95% confident that the true proportion of ninth grade math students that prefer to do their math work with pen is between 0.5393 and 0.6607.

Part (b)
The school paper states that more than half of ninth grade math students prefer to do their math in pen. 50% or 0.5 is below our confidence interval. Because our confidence interval contains values that are above 0.5, we believe that the newspaper has made an accurate statement.

Scoring

Parts (a) and (b) are essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is correct if each of the following is included:
Identify the appropriate confidence interval by name or formula
Check appropriate conditions
Correct mechanics
Interpret the confidence interval in context

Part (a) is partially correct if only three of the above are provided in the answer.

Part (a) is incorrect if less than three of components are provided in the answer.

Part (b) is essentially correct if the student connects the paper's claim of 0.5 to the confidence interval and says that the interval supports the paper's claim.

Part (b) is partially correct if the student says that the confidence interval does support the paper's claim, but have weak or incomplete explanation.

Part (b) is incorrect if the student fails to provide an explanation.

4 Complete Response

Both parts essentially correct.

3 Substantial Response

One part essentially correct and one part partially correct

2 Developing Response

One part essentially correct and one part incorrect

OR

Partially correct on both parts

1 Minimal Response

Partially correct on one part.

Free Response 2. **Solution**

Part (a)

1. Identify the population of interest and define the parameters of interest being estimated.

The population of interest is price of hotel rooms in a major city.

μ_{peak} = the population mean cost of a hotel room in peak season.

μ_{off} = the population mean cost of a hotel room in off season.

$\bar{x}_{peak} = 245$ = the sample mean cost of a hotel room in peak season.

$\bar{x}_{off} = 135$ = the sample mean cost of a hotel room in off season.

$s_{peak} = 45$ = sample standard deviation cost of a hotel room in peak season.

$s_{off} = 65$ = sample standard deviation cost of a hotel room in off season.

$n_{peak} = 33$ is the sample size of the cost of a hotel room in peak season.

$n_{off} = 38$ is the sample size of the cost of a hotel room in off season.

df = 33 – 1 = 32 degrees of freedom (using the simplified approximation for df).

2. Identify the appropriate confidence interval by name or formula.

We will use a 95% confidence t-interval for the difference of means (2-sample-t-interval).

3. Verify any conditions (assumptions) that need to be met for that inference procedure.

The problem states that each is a simple random sample.

The problem states that the assumption of normality is reasonable.

The population standard deviation is unknown.

4. Calculate the confidence interval.

At a 95% CI, the critical value is $t^* = 2.042$ from the table using df = 30.

This value is found on the t-distribution table using 30 degrees of freedom (because the table does not have a value for 32).

$$(245 - 135) \pm 2.042\sqrt{\frac{45^2}{33} + \frac{65^2}{38}}$$

$$110 \pm 26.82$$

$$(83.18, 136.82)$$

5. Interpret your results in the context of the situation.

We are 95% confident that the true difference in mean cost of a hotel room in this city between peak season and off season falls between \$83.18 and \$136.82.

Note: The calculator gives a result of (\$83.77, \$136.23) for this confidence interval. df = 65.9

Part (b)

The difference in cost from peak season to off season is 218-88 = \$130

Because \$130 falls inside our confidence interval, we would not think that this is an extreme change in price from peak to off season.

Scoring

Parts (a) and (b) are essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is correct if each of the following is included:

Identify the appropriate confidence interval by name or formula
Check appropriate conditions
Correct mechanics
Interpret the confidence interval in context

Part (a) is partially correct if only three of the above are provided in the answer.

Part (a) is incorrect if less than three of components are provided in the answer.

Part (b) is essentially correct if the student compares the difference in prices to the confidence interval and concludes that this is not an unusual price change for this city.

Part (b) is partially correct if the student says that this in not an unusual price change but does not directly tie their conclusion to the confidence interval.

Part (b) is incorrect if fails to provide an explanation.

4 Complete Response

Both parts essentially correct.

3 Substantial Response

One part essentially correct and one part partially correct

2 Developing Response

One part essentially correct and one part incorrect

OR

Partially correct on both parts

1 Minimal Response

Partially correct on one part.

AP Statistics Exam Connections

The list below identifies free response questions that have been previously asked on the topic of Confidence Intervals. These questions are available from the CollegeBoard and can be downloaded free of charge from AP Central http://apcentral.collegeboard.com.

Free Response Questions	
2002 B	Question 4
2004	Question 6
2005	Question 5
2006	Question 4