1-Sample *t* Confidence Intervals for Means

Requirements for complete responses to free response questions that require 1-sample *t* confidence intervals for means:

1. Identify the population parameter of interest, the type of inferential procedure to be used, and the confidence level.
   - A confidence interval is used to estimate the value of an unknown *parameter*.
   - The *confidence level* gives the percent of intervals produced by the procedure you used that contain the true value of the parameter.
   - The population parameter of interest must be defined in context of the problem.
2. State and verify whether the conditions (assumptions) needed for this procedure are met.
   - The conditions for a 1-sample *t*-interval are as follows:
     - The sample must be a simple random sample from the population of interest. Usually this is stated in the prompt. You must state this condition and verify whether or not this condition is met.
     - The sample must be large OR the population must be normally distributed. To verify this condition check that $n \geq 30$. If the sample size is small, you must show a graph of the sample data and state how the graph shows whether it is reasonable to assume the population is normally distributed or not.
     - The individual observations are independent. When sampling without replacement, the population must be at least 10 times greater than the sample size. $(N \geq 10n)$
3. Write the formula for the confidence interval and determine the interval.
   - Every confidence interval computation follows the same general form (given on the formula sheet):
     (Sample statistic) $\pm$ (distribution critical value) $\times$ (standard deviation of statistic)
   - The formula for computing a one-sample *t* confidence interval for means is:
     - $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$, where t^* is the critical value for $(n-1)$ degrees of freedom and the standard error of the mean is $\frac{s}{\sqrt{n}}$
4. Interpret the interval in the context of the problem.
   - Include a statement that links the *confidence level* to the interval.
   - When you interpret a confidence interval, be very careful that you do not associate the probability with the parameter. Keep in mind that the value of the parameter does not vary, so there is no chance associated with its value. The sample is what varies!
   - It is risky to try to interpret a confidence interval in a creative way. Learn a good confidence interval interpretation and stick to it! A suggested writing template for the interval is:

     We are _____% confident that the true mean "context of problem" is between _______and_________(units).

If you are asked to interpret the confidence level, the following is an appropriate interpretation of confidence level:

_____% of the intervals produced using this method will capture the true mean "context of problem."

Selecting a sample size:

When planning a study, the sample size chosen allows us to estimate a population mean within a given margin of error.

- The formula to determine the sample size based on the desired confidence level is $\text{ME} = t^* \frac{s}{\sqrt{n}}$. Note: you must assume $df = \infty$ to determine t^*.
- When you calculate the sample size needed to obtain a certain margin of error, the answer must be an integer equal to or greater than the value you obtain algebraically. (round to the greatest whole number)
- Increasing the confidence level results in a larger margin of error (if all else remains the same).
- Increasing the sample size results in a smaller margin of error (if all else remains the same).

Multiple Choice Questions:

1.

A test engineer wants to estimate the mean gas mileage μ (in miles per gallon) for a particular model of automobile. Eleven of these cars are subjected to a road test, and the gas mileage is computed for each car.

A dotplot of the 11 gas-mileage values is roughly symmetrical and has no outliers. The mean and standard deviation of these values are 25.5 and 3.01, respectively. Assuming that these 11 automobiles can be considered a simple random sample of cars of this model, which of the following is a correct statement?

A. A 95% confidence interval for μ is $25.5 \pm 2.228 \times \frac{3.01}{\sqrt{11}}$.

B. A 95% confidence interval for μ is $25.5 \pm 2.201 \times \frac{3.01}{\sqrt{11}}$.

C. A 95% confidence interval for μ is $25.5 \pm 2.228 \times \frac{3.01}{\sqrt{10}}$.

D. A 95% confidence interval for μ is $25.5 \pm 2.201 \times \frac{3.01}{\sqrt{10}}$.

E. The results cannot be trusted; the sample is too small.

2.

A random sample has been taken from a population. A statistician, using this sample, needs to decide whether to construct a 90 percent confidence interval for the population mean or a 95 percent confidence interval for the population mean. How will these intervals differ?

A. The 90 percent confidence interval will not be as wide as the 95 percent confidence interval.
B. The 90 percent confidence interval will be wider than the 95 percent confidence interval.
C. Which interval is wider will depend on how large the sample is.
D. Which interval is wider will depend on whether the sample is unbiased.
E. Which interval is wider will depend on whether a z-statistic or a t-statistic is used.

3.

A simple random sample produces a sample mean $\bar{x}$ of 15. A 95 percent confidence interval for the corresponding population mean is 15 ± 3. Which of the following statements must be true?

A. Ninety-five percent of the population measurements fall between 12 and 18.
B. Ninety-five percent of the sample measurements fall between 12 and 18.
C. If 100 samples were taken, 95 of the sample means would fall between 12 and 18.
D. $P(12 \leq \bar{x} \leq 18) = 0.95$
E. If $\mu = 19$, this $\bar{x}$ of 15 would be unlikely to occur.

4.

An engineer for the Allied Steel Company has the responsibility of estimating the mean carbon content of a particular day's steel output, using a random sample of 15 rods from that day's output. The actual population distribution of carbon content is not known to be normal, but graphic displays of the engineer's sample results indicate that the assumption of normality is not unreasonable. The process is newly developed, and there are no historical data on the variability of the process. In estimating this day's mean carbon content, the primary reason the engineer should use a t-confidence interval rather than a z-confidence interval is because the engineer

A. is estimating the population mean using the sample mean
B. is using the sample variance as an estimate of the population variance
C. is using data, rather than theory to judge that the carbon content is normal
D. is using data from a specific day only
E. has a small sample, and a z-confidence interval should never be used with a small sample

5.

A quality control inspector must verify whether a machine that packages snack foods is working correctly. The inspector will randomly select a sample of packages and weigh the amount of snack food in each. Assume that the weights of food in packages filled by this machine have a standard deviation of 0.30 ounce. An estimate of the mean amount of snack food in each package must be reported with 99.6 percent confidence and a margin of error of no more than 0.12 ounce. What would be the minimum sample size for the number of packages the inspector must select?

A. 8
B. 15
C. 25
D. 52
E. 60

6.

Which of the following statements is (are) true about the t-distribution with k degrees of freedom?

T I. The t-distribution is symmetric.
T II. The t-distribution with k degrees of freedom has a smaller variance than the t-distribution with $k+1$ degrees of freedom.
F III. The t-distribution has a larger variance than the standard normal (z) distribution.

A. I only
B. II only
C. III only
D. I and II
E. I and III

Free Response Questions:

7. An environmental group conducted a study to determine whether crows in a certain region were ingesting food containing unhealthy levels of lead. A biologist classified lead levels greater than 6.0 parts per million (ppm) as unhealthy. The lead levels of a random sample of 23 crows in the region were measured and recorded. The data are shown in the stemplot below.

Lead Levels

2	8
3	0
3	588
4	112
4	688
5	012234
5	99
6	34
6	68

Key: 2|8 = 2.8 ppm

(a) What proportion of crows in the sample had lead levels that are classified by the biologist as unhealthy?

(b) The mean lead level of the 23 crows in the sample was 4.90 ppm and the standard deviation was 1.12 ppm. Construct and interpret a 95 percent confidence interval for the mean lead level of crows in the region.

8. Biologists believe they have discovered a connection between foot size and back problems for gorillas in subtropical Africa. To study the size of the feet of gorillas in a large region where the back problems are common, they randomly selected 20 of the gorillas from the region and measured the length of their right foot. Some statistics resulting from this random sample are as follows.

Sample size	20	Minimum	11.2 cm
Mean	20.8 cm	First quartile	14.7 cm
Standard deviation	7.3 cm	Median	19.5 cm
		Third quartile	26.0 cm
		Maximum	33.0 cm

The biologists would like to construct a 95 percent confidence interval for the mean right foot length of the gorillas in the region.

(a) State and verify the conditions necessary in order for this confidence interval to be appropriate?

(b) The 95% interval constructed from this sample is (17.38, 24.22). Give an interpretation of the confidence interval and give an interpretation of the confidence level.

9. A car manufacturer is interested in conducting a study to estimate the mean stopping distance for a new type of brakes when used in a car that is traveling at 60 miles per hour. These new brakes will be installed on cars of the same model and the stopping distance will be observed. The cost of each observation is \$100. A budget of \$12,000 is available to conduct the study and the goal is to carry it out in the most economical way possible. Preliminary studies indicate that $\sigma = 12$ feet for the stopping distance.

   (a) Are sufficient funds available to estimate the mean stopping distance to within 2 feet of the true mean stopping distance with 95% confidence?

   Explain your answer.

   (b) A regulatory agency requires a 95% level of confidence for an estimate of mean stopping distance that is within 2 feet of the true mean stopping distance. The car manufacturer cannot exceed the budget of \$12,000 for the study. Discuss the consequences of these constraints.