In-class problems on confidence intervals

Decide whether the following statements are true or false.  Explain your reasoning.

Problems:

a)  For a given standard error, lower confidence levels produce wider confidence intervals.

b)  If you increase sample size, the width of confidence intervals will increase.

c)  The statement, "the 95% confidence interval for the population mean is (350, 400)", is equivalent to the statement, "there is a 95% probability that the population mean is between 350 and 400".

d)  To reduce the width of a confidence interval by a factor of two (i.e., in half), you have to quadruple the sample size.

e)  Assuming the central limit theorem applies, confidence intervals are always valid.

f)  The statement, "the 95% confidence interval for the population mean is (350, 400)" means that 95% of the population values are between 350 and 400.

g)  If you take large random samples over and over again from the same population, and make 95% confidence intervals for the population average, about 95% of the intervals should contain the population average.

h)  If you take large random samples over and over again from the same population, and make 95% confidence intervals for the population average, about 95% of the intervals should contain the sample average.

i)   It is necessary that the distribution of the variable of interest follows a normal curve.

j)  A 95% confidence interval obtained from a random sample of 1000 people has a better chance of containing the population percentage than a 95% confidence interval obtained from a random sample of 500 people.

k)  If you make go through life making 99% confidence intervals for all sorts of population means, about 1% of the time the intervals won't cover their respective population means.

Problems:

a)  For a given standard error, lower confidence levels produce wider confidence intervals.

False.   To get higher confidence, we need to make the interval wider interval.  This is evident in the multiplier, which increases with confidence level.

b)  If you increase sample size, the width of confidence intervals will increase.

False.   Increasing the sample size decreases the width of confidence intervals, because it decreases the standard error.

c)  The statement, "the 95% confidence interval for the population mean is (350, 400)", is equivalent to the statement, "there is a 95% probability that the population mean is between 350 and 400".

False.   95% confidence means that we used a procedure that works 95% of the time to get this interval.  That is, 95% of all intervals produced by the procedure will contain their corresponding parameters.  For any one particular interval, the true population percentage is either inside the interval or outside  the interval.  In this case, it is either in between 350 and 400, or it is not in between 350 and 400.  Hence, the probabliity that the population percentage is in between those two exact numbers is either zero or one.

d)  To reduce the width of a confidence interval by a factor of two (i.e., in half), you have to quadruple the sample size.

True, as long as we're talking about a CI for a population percentage.   The standard error for a population percentage has the square root of  the sample size in the denominator.  Hence, increasing the sample size by a factor of 4 (i.e., multiplying it by 4) is equivalent to multiplying the standard error by 1/2.  Hence, the interval will be half as wide.  This also works approximately for population averages as long as the multiplier from the t-curve doesn't change much when increasing the sample size (which it won't if the original sample size is large).

e)  Assuming the central limit theorem applies, confidence intervals are always valid.

By "valid," we mean that the confidence interval procedure has a 95% chance of producing an interval that contains the population parameter.

False.  The central limit theorem is needed for confidence intervals to be valid.   However, it is also necessary that the data be collected from random samples.  Confidence intervals will not remedy poorly collected data.

f)  The statement, "the 95% confidence interval for the population mean is (350, 400)" means that 95% of the population values are between 350 and 400.

False.  The confidence interval is a range of plausible values for the population average.   It does not provide a range for 95% of the data values from the population.  To find the percentage of values in the population between 350 and 400, we need to look at a histogram of the data values and determine what percentage of observations are between 350 and 400.

g)  If you take large random samples over and over again from the same population, and make 95% confidence intervals for the population average, about 95% of the intervals should contain the population average.

True.   This is the definition of confidence intervals.

h)  If you take large random samples over and over again from the same population, and make 95% confidence intervals for the population average, about 95% of the intervals should contain the sample average.

False.   The confidence interval is a range for the population average, not for the sample average.  In fact, every confidence interval contains its corresponding sample average, because CIs are of the form:  sample avg. +/- multiplier SE.  So, the sample average is right in the middle of the CI.

i)   It is necessary that the distribution of the variable of interest follows a normal curve.

False.   It is necessary that the distribution of the sample average follows a normal curve.  The data values of the variable, however, need not follow a normal curve, because if the sample size is large enough the central limit theorem for the sample average will apply.

j)  A 95% confidence interval obtained from a random sample of 1000 people has a better chance of containing the population percentage than a 95% confidence interval obtained from a random sample of 500 people.

False.  All 95% confidence intervals have the property that they come from a procedure that has a 95% chance of yielding an interval that contains the true value.   The confidence interval method automatically accounts for sample size in the standard error.   A 95% CI with n=1000 will be narrower than a 95% CI with n=500, but both CIs will have 95% confidence of containing the population percentage.

k)  If you make go through life making 99% confidence intervals for all sorts of population means, about 1% of the time the intervals won't cover their respective population means.

True.  Since 99% of the intervals should contain the corresponding population mean, 1% of them will not.