Bonnie Turek

Eco 602 – Week 7 Reading Questions

10/12/21

Confidence Intervals

For Questions 1 - 4, assume you are working with a population that is normally-distributed with mean μ and standard deviation σ. Note that although these population parameters exist, you cannot know their exact values and you must estimate them through sampling.

*The width of confidence intervals is influenced by properties of both the population and the sampling process. *Recall that we are not 95% certain that a 95% confidence interval we calculate contains the true value.

Q1. Explain the effect, if any, of the population mean on the width of CIs for a population that is normally distributed. If population mean does not affect the widths of CIs explain why not.

NO, the population mean would not affect the width of confidence intervals. Typically, we are using confidence intervals to estimate a range of values for an unknown population parameter, like the population mean. We do not know the population mean, because an assumption of Frequentist statistics is that populations are large and unknowable, but there is a true value. That is why we use samples from a population to make inferences about a population, like what is the population mean for example. Therefore, I would say that whatever the unknown population mean is, must inherently affect the confidence interval **values**, but changing the (unknown) population mean would not change the **width** of the intervals. The population is normally distributed, so if the population mean were to change, it would just shift the normal distribution curve and alter confidence interval values, not the widths. It is almost the other way around, in that the confidence interval width tells us more or less precise information about the possible population mean values.

Q2. Explain the effect, if any, of the population standard deviation on the width of CIs. If population standard deviation does not affect the widths of CIs explain why not.

YES, Population standard deviation does affect the width of confidence intervals. The population standard deviation refers to the measure of the amount of variation or dispersion of a set of values in our population. In calculating confidence intervals, it would be best to use the population standard deviation, but we often do not know it. Therefore, we can use standard deviation of many repeated samples. This should give us an idea of the possible amount of dispersion amongst sample and population values. Typically, greater standard deviation values and therefore, greater variance and dispersion in our observations, yields greater uncertainty or precision in our statistical inferences of the population. Overall, you would assume the greater the population standard deviation, the wider the confidence intervals will be. The width increases as the standard deviation increases. Wider intervals mean less precise inferences about the population.

Q3. Explain the effect, if any, of the population size on the width of CIs. If population size does not affect the widths of CIs explain why not.

NO, Population size should not affect the width of confidence intervals. Unless, however, you are dealing with a very, very small population. This is typically not the case, since Frequentist statistics assume the population is large and infinite. Often you may not even know the exact population size. Since population size is irrelevant for confidence interval widths, a sample of 500 people is equally useful in examining the opinions of a state of population size 15,000,000 as it would a city of population size 100,000. However, it is important to recognize that our **sample** size does influence the width of the confidence interval. Overall, we can ignore the population size when it is “large” or unknown.

Q4. Explain the effect, if any, of the sample size on the width of CIs. If sample size does not affect the widths of CIs explain why not.

YES, Sample size would affect the width of confidence intervals. Sample size is referring to the number of observations or replicates to include in a statistical sample. It is important to choose a good sample size in order to make meaningful inferences about your population in the study. Increasing sample size generally leads to increased precision when estimating unknown population parameters. Therefore, increasing sample size decreases the width of confidence intervals, because it decreases the standard error of the sample, and allows for increased precision. (Standard error refers to the variance of the sample means and the error in using the sample mean to make inferences about the population mean.) Generally, wider confidence intervals are less precise, and narrower confidence intervals are more precise.

Q5. Interpreting a CI. Use a narrative example of a real (or made up) dataset to describe what a Frequentist 95% confidence interval really means. Make sure you cover any relevant assumptions of the Frequentist paradigm, and your answer must be in non-technical language.

A frequentist 95% confidence interval is constructed such that if the model assumptions are correct, and under hypothetical repeated sampling, 95% of the intervals constructed would contain the true value of the parameter. As an example, suppose you are working with the height measurements of 40 randomly chosen women. Our sample is the 40 random women, and the population in this case is all women in the U.S. (or in the world, but let’s keep it limited to the U.S.). We don’t know the actual population mean of all women’s heights, so we are using samples to construct a confidence interval that will tell us some estimated information about the rest of the population. Suppose our sample of 40 women yields a mean height of 165 cm and a standard deviation of 15 cm. We can use those values to construct a 95% confidence interval, such that our result says the true mean of ALL women (if we could measure all their heights) is likely to be between 158.8cm and 171.2cm. The confidence interval range here as an example was 165 +/- 6.2 cm. So, 165 + 6.2 = 171.2cm, the upper limit of the interval, and 165 – 6.2 = 158.8cm, the lower limit of the confidence interval. The "95%" says that with many repeated experiments like we just did, measuring 40 random women’s heights, we expect that the true mean height of women would fall within my 95% confidence intervals approximately 95% of the time. So, there is a 1-in-20 chance (5%) that our Confidence Interval does NOT include the true mean of height.