1. Introduction
   1. My name is Matthew Massey and this is the video portion of our midterm exam for Statistics 570…Summer semester of 2021…
   2. There are two parts which I will answer over the next few slides…
   3. Click
2. The first part asks the question…Why do we not usually use 99.99% confidence intervals?
   1. First of all…a Confidence interval tells us how good a point estimate is…or how good our procedure predicts the parameter
   2. For example… “We are 95% confident that the range between X and X+ contains the true population parameter.”
   3. We can use different confidence levels to construct these intervals…some of the common ones are 90%, 95%, and 99%...the three graphs on the bottom show a standard normal distribution with these three confidence levels shaded in red…
   4. You’ll notice that the higher the confidence level…the wider the confidence interval…
   5. Although the higher confidence level of 99% shows a wider range of values…there is also less risk of missing the true value of the population parameter
   6. click
   7. With a 99.99% confidence level…shown here on the right…our confidence interval becomes even wider and it would be extremely rare for a value to NOT be in this range…so it would be hard to distinguish a good estimate of the parameter from a bad estimate
   8. There’s also very little difference in risk of MISSING the population parameter between the 99% and 99.99% confidence intervals…so not much to gain for most applications
3. The second consists of several questions based on this information…Suppose that UK’s Alumni Office wants to collect sample data to investigate whether UK graduates from the math department differ from UK graduates from the mechanical engineering department with regard to average annual salary….and now I’ll go through the associated questions…
   1. What are the observational units?
      1. In this observational study…we can define the units that will be measured…or observational units…as the annual salaries of graduates in the math and mechanical engineering departments
   2. What is the response variable?
      1. This study aims to find out if there is a relationship between annual salaries of recent graduates versus their respective major department…math or mechanical engineering….So the response variable is annual salary…which may be explained by the graduation department.
   3. Should you advise the alumni office to use random sampling to collect the data?
      1. Yes…the logic behind random sampling is to use chance as a way to create two groups that are similar. For a simple random sample we would need to choose a sample of n subjects in such a way that every other possible sample of the same size n has the same chance of being chosen…
   4. Should you advise the alumni office to use random assignment to collect the data?
      1. Random assignment refers to a method used to place participants into groups. This is already predetermined by each graduate’s major department…math or mechanical engineering…so NO there is no need for random assignment in this study
   5. What is the alternative hypothesis (or research question) to be tested? Describe in words (not symbols).
      1. The study is designed to compare salaries of graduates from the math department versus the mechanical engineering department.
      2. In other words, the study will test the alternative hypothesis that the mean salary of graduates from the math department is different than the mean salary of graduates from the mechanical engineering department.