***Answer each question below on separate sheets of paper. Make sure to clearly label each of your answers (e.g., #1 or #4a), put your name on each extra sheet used, and staple these questions to the top of your answer sheets when completed to hand in.***

**library(NCStats)**

**distrib(val,mean=meanval,sd=sdval,lower.tail=FALSE,type=”q”)**

where

* **val** is a value of the quant. variable or area (i.e., percentage as a proportion)
* **meanval** is population mean ()
* **sdval** is standard deviation () or error (SE)
* **lower.tail=FALSE** is included for “right-of” calculations
* **type=”q”** is included for reverse calculations

1. **[5 pts]** Completely describe three major differences between a population distribution and a sampling distribution.
2. **[20 pts]** A company produces packages of grape tomatoes that are advertised as weighing 16 ounces. Suppose that it is known that the distribution of package weights is slightly right skewed with a mean of 16.1 ounces and a standard deviation of 0.3 ounces. Use this information to answer the questions below. *If you choose not to answer the question, carefully and specifically explain why. If you choose to answer the question, please write your final numerical answer (to one decimal place) along with the R code that you used.*
3. What is the probability that a package weighs more than 17 ounces?
4. What is the probability that the mean weight of 20 packages will be less than 16 ounces?
5. What is the probability that the mean weight of 40 packages will be more than 16 ounces?
6. What is the probability that the mean weight of 25 packages will be between 16.0 and 16.2 ounces?
7. What is the IQR for the mean weights of 20 packages?
8. **[5 pts]** What is the definition of a p-value and how is it used to make a decision about H0?
9. For the following questions, use HA: >500, =150, n=25, and =0.10. *Show your work and R code where appropriate.*
   1. **[5 pts]** Compute the p-value ifx=530.
   2. **[2 pts]** Make a decision about HO from your computed p-value.
   3. **[5 pts]** Compute an appropriate confidence region ifx=530.
   4. **[2 pts]** Interpret your confidence region.
   5. **[5 pts\*]** Compute  assuming that the actual  is 520.
10. A dean wants to determine if it takes more than 10 minutes, on average, for students to get from one class to another. In an effort to test this hypothesis, she collected a random sample of 100 between-class travel times and found the mean to be 10.22 mins. Assume that it is known from previous studies that the distribution of between-class travel times is symmetric with a standard deviation of 1.60 mins. Test the dean’s hypothesis with α=0.10. Use this information to answer the questions below. *Show your work and R code where appropriate.*
    1. **[6 pts]** What are the null and alternative hypotheses? [*Make sure to define the parameter in the hypotheses.*]
    2. **[5 pts]** Compute the p-value.
    3. **[3 pts]** Make a decision; i.e., what can be concluded about the time it takes for students to move between classes? [*use a complete sentence*]
    4. **[5 pts]** Construct an appropriate confidence region.
    5. **[3 pts]** Interpret your confidence region. [*use a complete sentence*]
    6. **[4 pts]** Define what a Type I and a Type II error would be in this study. [*use complete sentences*]
11. **[5 pts]** Foresters want to sample enough trees from a large stand of trees such that the mean diameter-at-breast-height (DBH) for all trees in the stand will be estimated to within 10 cm with 99% confidence. Suppose from a pilot study sample of a few trees, that the researchers estimated the standard deviation of DBH among trees to be 60 cm. Use this information to determine how many trees should be sampled to meet the forester’s constraints. *Show your work and R code where appropriate.*
12. **[6 pts]** Suppose that the following sets of sample means were extracted from a population with a population mean of 40.

Set #1. – 36, 37, 38, 39

Set #2. – 38, 38, 41, 43

Set #3. – 10.2, 20.8, 30.4, 40.5

Set #4. – 15.1, 30.3, 39.7, 74.9

Identify which set from above **best** represents each situation below.

* 1. Accurate and most precise.
  2. Inaccurate and most precise.
  3. Inaccurate and least precise.

1. **[5 pts]** Describe choices that you, as a researcher, can make to increase statistical power. Which is the best choice to make (statistically) and why?
2. **[5 pts]** Describe choices that you, as a researcher, can make to reduce the margin-of-error. Which is the best choice to make (statistically) and why?