ANSC/TPSS 603
Assignment 1 - Review

This assignment is based on standard material from an introductory statistics course, and should be a straight-forward review for you. It is intended to help you check you preparation for this course. Don't worry if you find that you are a little rusty on some aspects of the material, and need to turn to your old lecture notes or texts for help. You may not even have covered some topics, so make a note to yourself of which questions are not on familiar material. (If almost nothing looks familiar, however, find out why.)

1. Given a data set of values X (milk quality measurements):

   1, 10, 7, 5, 12, 9, 16

   a. Find the mean, median, variance, s^2, standard deviation, s, and coefficient of variation, CV. Determine the degrees of freedom, df, of s^2 and of s and find the standard error of the mean.

   b. Find the mean, median, variance, standard deviation, and coefficient of variation, degrees of freedom of s^2 and s, and standard error of the mean for the same data set after you add 15 to each data value. Show how each formula is changed.

   c. Repeat, but this time multiply each data value by 4, rather than adding 15. Show how each formula is changed.

2. If a data set has n = 7 observations with $\sum X = 60$ and $\sum X^2 = 656$, calculate as many of the mean, median, variance, standard deviation, and coefficient of variation as you can, given only that information.

3. If t follows the student's t distribution with 6 degrees of freedom (df):

   a. Find the probability that absolute value of t is at least 3.707, i.e. $P(|t| \geq 3.707)$. What would this probability be, roughly, if the df were 40?

   b. Find the value of t_o which is larger than the absolute value of t with 95% probability, i.e. so that $P(|t| \leq t_o) = 0.95$. What would t_o be if the df were 40?

4. a. Construct a 95% confidence interval for the unknown mean of the population from which the sample of 7 observations X in question 1 was drawn, the individual milk yields of several hundred dairy cows. Assume the sample was taken randomly.

   b. How would the population represented and reliability of your answer be affected if the sample was not random, but rather was selected:

      1. By an expert in the field of testing milk quality?

      2. From seven cows in one randomly selected herd?

      3. From the seven best cows in a previous testing?