STAT 477/STAT 577
HW 1 - Module 1: Sections 1 and 2

1. In lecture, we discussed a course survey that was given to students enrolled in STAT 101 during the Fall 2014 semester. Links to the survey and the data collected can be found in the **OpeningSurvey.pdf** and **OpeningSurveyData.csv** files in Canvas. Select three categorical variables from the survey and use R to summarize each variable with its summary table and bar graph.

2. A certain genetic mutation occurs in a population with probability 0.05. A researcher has genetic material from 40 unrelated members of this population and tests for the mutation.

   (a) The number of people in a sample of 40 unrelated members of this population with this genetic mutation has a binomial distribution. What are the values of the parameters for this binomial distribution (n and p)?

   (b) Use R to calculate the probability that at least 1 person in a sample of 40 unrelated members of this population will have the genetic mutation.

   (c) Use R to calculate the probability that no more than 3 people in a sample of 40 unrelated members of this population will have the genetic mutation.

   (d) What is the mean number of people with the genetic mutation in a sample of 40 unrelated members of this population?

   (e) What is the variance and standard deviation of the number of people with the genetic mutation in a sample of 40 unrelated members of this population?

   (f) Use R to produce a graph of the distribution of the number of people with the genetic mutation in a sample of 40 unrelated members of this population. Describe the shape of the distribution.

3. Cocker spaniels (a breed of dog) are susceptible to anemia. Suppose that 30% of the population of seven year old cocker spaniels have anemia.

   (a) The number of cocker spaniels with anemia in a sample of 40 dogs from this population has a binomial distribution. What are the values of the parameters for this binomial distribution (n and p)?

   (b) Use R to calculate the probability that at least 13 of the dogs in a sample of 40 dogs from this population will have anemia.

   (c) Use R to calculate the probability that no more than 8 dogs in a sample of 40 dogs from this population will have anemia.

   (d) What is the mean number of dogs with anemia in a sample of 40 dogs from this population?

   (e) What is the variance and standard deviation of the number of dogs with anemia in a sample of 40 dogs from this population?

(f) Use R to produce a graph of the distribution of the number of dogs with anemia in a sample of 40 dogs from this population. Describe the shape of the distribution.

4. Suppose, based on numerous chess games between these two players, it has been determined the probability Player A would win is 0.40, the probability Player B would win is 0.35, and the probability the game would end in a draw is 0.25.

   (a) Find the probability that Player A would win 7 games, Player B would win 2 games, and the remaining 3 games would each end in a draw if they played 12 games.

   (b) Find the expected number of games Player A would win and the expected number of games Player B would win if the two players played 12 games.

   (c) Find the variance of the number of games Player A would win and the variance of the number of games Player B would win if the two players played 12 games.

   (d) Find the correlation of the number of games won between Player A and Player B.