Instructions: Submit one .pdf file that includes the answers to both sets of questions
Please feel free to discuss questions on the discussion board.

Computational Question

1. (5 points) This question uses the `gifted` dataset, part of the `openintro` R package. You will need to install the package (do this once and don't store it in your R script):

   ```
   install.packages("openintro")
   ```

   Then load the package (this should be in your R script)

   ```
   library("openintro")
   ```

   You will then have access to the `gifted` data frame

   ```
   gifted       # print in its entirety
   head(gifted) # the first six rows
   ?gifted      # read the documentation for this dataset
   ```

   The column `motheriq` contains the mother's IQ for 36 gifted children. We are interested in whether the mothers of gifted children have an IQ higher than the population at large, which is 100.

   (a) State the null and alternative hypothesis in statistical notation, and in words.

   (b) Give the formula for the test statistic you will use, and calculate it.

   (c) Give the p-value for the test, and the line of code you used to calculate it.

   (d) Calculate a point estimate and a 95% confidence interval for the mean IQ of mothers of gifted children.

   (e) Give a summary of your findings.

Conceptual Questions

2. (2 points) In your own words, describe how understanding the behavior of a statistic over many samples (the sampling distribution) allows us to make inference about a population from a single sample.

   Answer any three of the following four short answer questions.

3. (1 point) A game show selects participants randomly from four regions in the United States: The northwest, the southwest, the south, and the northeast. On each episode, one participant from each region competes. To date, this game show has selected 1000 participants, with 250 from each region. Thus, to date, there have been 250 games. Participants from the south have won 100 times.

   (a) Explain the general process in which we use probability to determine if participants from the South are more likely to win. That is, the population proportion of wins by participants from the south is greater than 0.25.

   (b) If we use the sample proportion to estimate the true proportion, what do we need to know to decide if the estimate unbiased?

(c) What will happen to the uncertainty of the estimate as the sample size increases? Intuitively, why is this the case?

4. (1 point) A random sample of 10 college graduates is selected to play Sudoku, with each graduate attempting the same puzzle. The mean time it takes them to correctly complete the puzzle is 5 min.
   (a) Which number should we use to estimate the population mean time it takes college graduates to finish this puzzle? Why?
   (b) Why is there uncertainty in this estimate?
   (c) What happens to the uncertainty if we increase the sample size to 1000 graduates?

5. (1 point) A farmer in Idaho estimates that the mean time it takes to grow one potato is 100 days. His 95% confidence interval is (90 days, 110 days).
   (a) What is wrong with the following interpretation of this confidence interval? *There is a 95% chance that the mean time it takes to grow one potato is between 90 and 110 days.*
   (b) Describe one correct way to interpret this confidence interval.
   (c) Describe what this confidence interval means in terms of probability.
   (d) How would the confidence interval change if the confidence level changes?

6. (1 point) A local restaurant is worried its potato supplier is skimping on the bags of potatoes it supplies. They claim each bag weighs 15lbs. The restaurateur understands that its hard to get a whole number of potatoes to weigh exactly 15lbs, but they feel it should average out. Over a month they weigh every bag they receive (and assume this is a random sample from all bags). They find a 95% confidence interval of the mean weight is 14.5 lbs to 14.9 lbs.
   (a) Would they reject or fail to reject the null hypothesis (at the 5% level) that the mean weight is 15lbs?
   (b) What can you say about the p-value from the hypothesis test?
   (c) What can you say about the sample mean from the measured bags?
   (d) What can you say about the p-value from the hypothesis test where the null hypothesis is the mean weight is 14.8 lbs?
   (e) The restaurateur presents this evidence to the supplier and asks for a refund. If you were the supplier, are there any problems you would bring up that cast doubt on the analysis?