1. Normal table: According to a national organization, the weights of six-year-old boys are normally distributed with mean 45 pounds and standard deviation 7 pounds.
   (a) What proportion of the population of six-year-old boys have weights between 40 and 60 pounds?
   (b) What weight is required to put someone in the top 15% of the distribution?

2. A company that manufactures light bulbs keeps a careful eye on the quality of their manufacturing process. They have determined that an appropriate rate of defective bulb production is 1% (or less). They measure the number of defective light bulbs in randomly selected boxes (each box contains 144 bulbs). Historically, the mean number of defectives is 1.4 (a bit less than 1% of the bulbs produced) and the standard deviation is 1.2.
   (a) Identify the population parameter of interest in this setting.
   (b) Explain why the number of defectives in a box would not be well described by a normal distribution.
   (c) Each week they sample 36 boxes and record $M$ the average number of defectives in the 36 sample boxes. Use the central limit theorem to identify the distribution of $M$.
   (d) One week the mean number of defectives $M$ is equal to 1.7 (approximately 1.2%). The quality control manager tells the president they should stop production and check the process out carefully.
      i. How likely is it to observe a value of $M$ bigger than or equal to 1.7?
      ii. Do you agree with the quality control manager?

3. Experiment or observational study? – A study was carried out to determine the effect of receiving free milk on the achievement of children in school. The idea is that many children don’t have adequate breakfast and it might be worthwhile to give free milk. The children in one school were randomly assigned by investigators into a milk group and a control group (no milk). After assignment, some teachers switched the assignments of some children to insure that especially needy children got milk. At the end of the year, reading levels were assessed for all children.
   (a) Identify the experimental units, treatments, and response variable for this study.
   (b) Is this study an experiment or observational study? Explain.
   (c) Explain why it was good for investigators to randomly assign students to the milk/control groups.
   (d) There are at least three ways to analyze the data from this experiment.
      i. The ”intention to treat” approach would compare the reading levels of the students that were initially randomly assigned to the milk group to those assigned to the control group. What are the advantages and disadvantages of this approach?
      ii. The ”as treated” approach would compare the reading levels of the students who ended up receiving milk to those that ended up not receiving milk. What are the advantages/disadvantages of this approach?
      iii. The ”per protocol” approach would exclude data from the children who got switched and analyze the data only for those children who stayed in the group they were assigned. What are the advantages/disadvantages of this approach?

4. $p$-values – The $p$-value is a central concept for significance testing in statistics. It is often misunderstood and has recently been identified as a key contributor to the lack of reproducibility of scientific studies (it appears that the results in a non-trivial fraction of scientific studies can not be reproduced). Read "The ASA’s Statement on p-Values" (available at course website or online). Should $p$-values be banned? Discuss positive and negative features of the use of $p$-values.

5. Install R and RStudio on your computer. See instructions on course website.