A quick trip through the text book shows that during our work in chapters 1, 3 and 4 that we have discussed the following:

- types of variables: categorical, quantitative;

- center, spread, shape

- stemplot

- histogram

- boxplot

- density plot

- modes,

- For unimodal distributions: symmetry, skew

- outliers (suspected outliers 1.5·IQR)

- mean, median

- quartiles, IQR, standard deviation

- five-number summary

- normal: rules of thumb, z-scores, standard normal, qqplot

- types of experiments

- comparative experiment

- placebo effect

- bias: lack of realism

- randomization

- randomized comparative experiment

- double blind, completely randomized

- matched pairs design

- population; sample

- statistic; parameter

- samples, SRS

- bias: response bias, wording of questions, lack of realism, ...

- sampling variability (Margin of Error)

- sampling distribution: center, spread, shape

- sample space, outcome, event, random variable

- rules of probability

- equally likely probabilities

- $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

- independence

- distribution of a discrete random variable X: $P(X = k) = ...$

- mean of a discrete random variable. Expectation, $E(X) = \mu = \Sigma k P(X = k)$

- variance and standard deviation of a discrete random variable. $\sigma^2 = \Sigma (k - \mu)^2 P(X = k)$.

Here are some sample problems to work on during class. These are not meant to exhaust the full range of questions I may ask.

1. A teacher has 4 items from the snack machine: a candy bar, a pack of gum, a bag of chips and a pack of licorice. She hands out 2 of them.

   (a) What is the probability the second one handed out is the pack of licorice?
   (b) What is the probability the last one handed out is the pack of licorice?

2. A randomly chosen student has a 30% chance of being enrolled in a math class and a 35% chance of being enrolled in an English class. Suppose that the chance they are enrolled in both a math class *and* an English clash is only 10%. What is the probability that a randomly chosen student is enrolled in either a math class *or* an English class?

3. In the NY Times (September 21, 2009) we find:

   > In a recent study for The Archives of Internal Medicine, scientists followed 153 men and women for two weeks, keeping track of their quality and duration of sleep. Then, during a five-day period, they quarantined the subjects and exposed them to cold viruses. Those who slept an average of fewer than seven hours a night, it turned out, were three times as likely to get sick as those who averaged at least eight hours.

   - Was this a comparative study?
   - Was this a double-blind study?
   - What is the "treatment" in the study?
   - Were the subjects randomized?

- Are there any issues in this study that you would think could introduce a bias?

4. In the NY Times (on 9/22/09) an article about flu shots reads:

> The latest trials, done in about 600 children and teenagers, ages 10 to 17, found that they received a âĂIJrobust immune responseâĂİ from a single dose, Dr. Fauci said. (Specifically, after 10 days, 76 percent had enough antibodies in their blood to be considered immune to swine flu. While that is not 100 percent protection, some could be expected to develop more antibodies after 10 days, and 76 percent is considered âĂIJnot bad at allâĂİ for flu vaccine, Dr. Fauci added.)
> . . .
> There were no serious reactions to the vaccine among the 600 children tested, Dr. Fauci said. Some complained of sore arms or mild fevers, he said, but that is typical for flu shots.

- Identify the population and the sample
- Is the 76% a population parameter or sample statistic in this article?
- If you found out that the sample was not chosen randomly, what would you say about the results?
- What is 600 in this story? (Use precise language)

5. If one tosses a fair coin 5 times a sequence results, say x_1, x_2, x_3, x_4, x_5, where each x is either H or T.

  (a) Find the probability of a sequence that produces 3 heads and 2 tails? (Does it matter which order the heads or tails are?)

  (b) There are 10 different sequences that will produce 3 heads and 2 tails. Each of these outcomes is disjoint from the others. Use the last question to find the probability one will get 3 heads in 5 coin tosses.

  (c) Explain, without doing any calculations, why this should be the same probability as getting 2 heads in 5 coin tosses *if* the coin is fair.

6. Recalling the rules of thumb for a normal, if math SAT scores for CSI students are normally distributed with a mean of 500 and standard deviation of 100, find the probability a randomly chosen student has an SAT score exceeding 600? Between 500 and 700?

7. Imagine that heights of 14-year old boys are normally distributed with a mean of 62 inches and a 4 inch standard deviation. If you pick a 14-year old boy at random

  (a) What is the probability they are 65 inches or taller?

  (b) What is the probability they are between 62 and 64 inches tall?

    (c) What height is such that there is a 25% the boy is less than this height.

8. A data set consists of values 1,1,2,3,4,5,8,15. Use the formulas or your calculator to find

    (a) What is n

    (b) What is $\bar{x}$?

    (c) What is s

    (d) What is the median?

    (e) What is the IQR

9. A boxplot of a data set is shown. Based on this identify

    (a) the median

    (b) The IQR

    (c) The max.

    (d) Will the mean or median be greater? Why

Benford's law?

sample(1:9, 100, replace = T, prob = p)
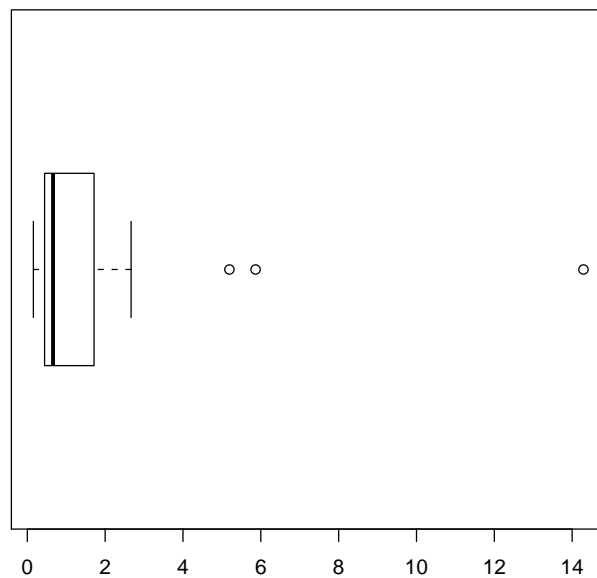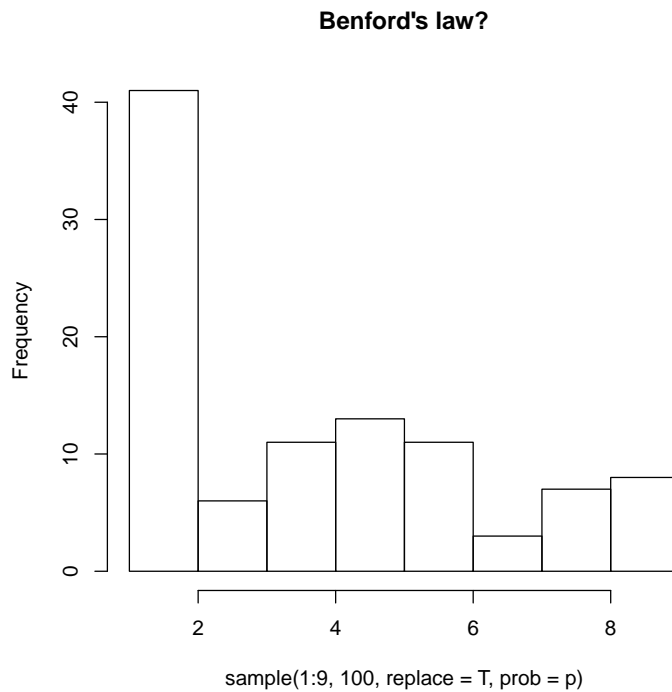
10.

A statistics teacher scrupolously records the first digit of each of receipts they receive. A histogram is shown.

(a) Is this data set unimodal, bimodal, multimodal?

(b) If unimodal, is this data set skewed right? skewed left? or symmetric?

(c) Estimate the mean value for this data set.

(d) There is something funny about this graphic in terms of representing the data. Can you tell what it is?

11. Which of these data sets appears to be normally distributed? Explain why.

d1

d2

d3

x
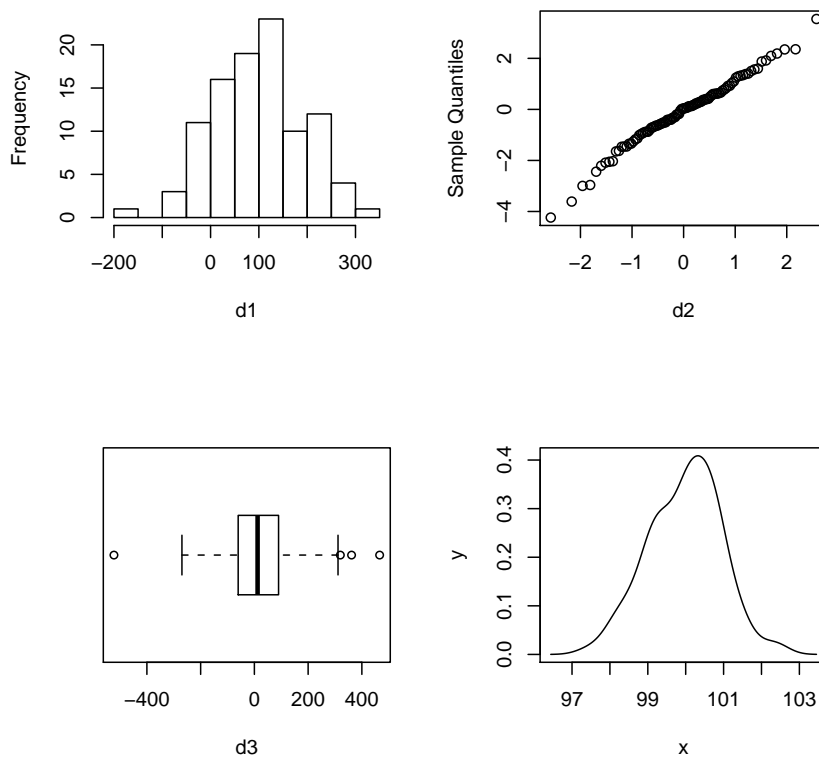
12. A 6-sided die has faces 1,1,2,3,5,8. Let X be the amount on one roll and Y the amount on 2 rolls or the die.

   (a) Write down the probability distribution for X

   (b) What is the expected value of X?

   (c) What is the expected value of Y?

   (d) The variance of X is $56/9$. What is the variance of Y?