Homework Assignment 1
(Due Monday, February 1, 2010)

1. Suppose a test for HIV correctly gives a positive result, if a person is infected, with probability 99.5%, and correctly gives a negative result, if a person is not infected, with probability 98%.

   (a) Suppose that 0.1% of a population are infected with HIV. Consider drawing a person at random and testing him or her for HIV infection. Calculate Pr(infected | test is positive).

   (b) Consider a person drawn from a high-risk group, so that they have, *a priori*, probability 30% of being infected. Calculate Pr(infected | test is positive).

   Do the calculations using two different approaches: 2×2 tables as in class, and Bayes' rule.

2. The seed of Mendel's pea plants were either smooth or wrinkled, the result of a single gene with two alleles, A (smooth) and a (wrinkled), with A dominant to a, so that seed with genotype AA or Aa are smooth and seed with genotype aa are wrinkled. $\rightarrow$ We cross two pure-breeding lines, one with smooth seed and genotype AA and the other with wrinkled seed and genotype aa, to obtain the F_1 with genotype Aa. $\rightarrow$ We self an F_1 and pick a random F_2 seed. $\rightarrow$ We grow up the F_2 and self it to obtain a random F_3 seed. Calculate the following:

   (a) $\Pr(F_2$ seed is smooth)

   (b) $\Pr(F_2$ seed has genotype Aa)

   (c) $\Pr(F_2$ seed has genotype Aa | it is smooth)

   (d) $\Pr(F_3$ seed is smooth | F_2 has genotype AA)

   (e) $\Pr(F_3$ seed is smooth | F_2 has genotype Aa)

   (f) $\Pr(F_3$ seed is smooth)

   (g) $\Pr(F_3$ seed is smooth | F_2 is smooth)

3. Consider the following data:

   ```
   12.5   10.0   9.2   10.0   9.1   11.0   10.7   9.2   10.1   9.1
   ```

   Calculate (on a piece of paper or with a hand calculator) the arithmetic mean, median, and SD of the above set of numbers. How would the mean, median, and SD change if you

   (a) added 2 to each data point?

   (b) multiplied each data point by 10?

   (c) multiplied each data point by -10?

   (d) added 2 to each data point and then multiplied each by 10?

   (e) multiplied each data point by 10 and then added 2 to each?
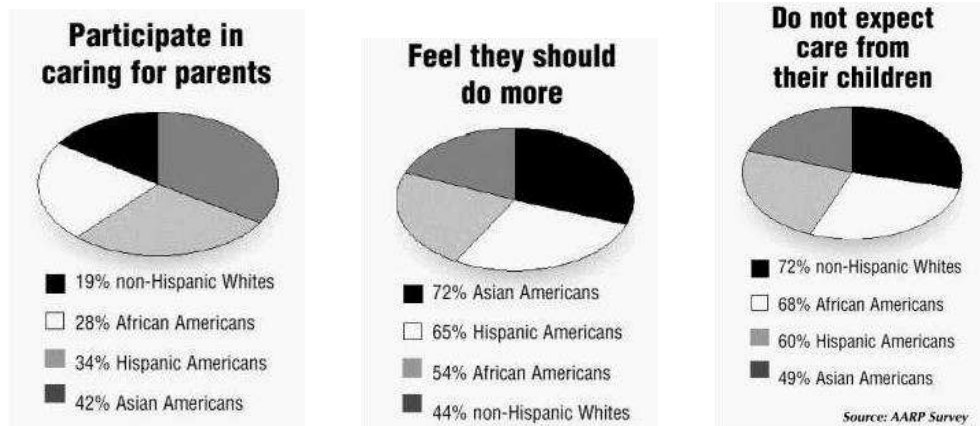
4. Consider the following histogram.
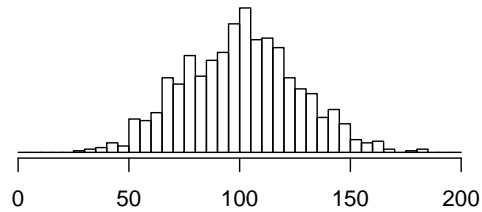
    (a) Which of the following is true? Choose one and explain briefly.

        i. The mean is larger than the median.

        ii. The median is larger than the mean.

        iii. The mean and the median are about the same.

    (b) Is the SD about 10, about 25 or about 50? Explain briefly.

5. Consider the following figures (taken from The Milwaukee Journal Sentinel, Friday, 18 Jan 2002, Senior Focus, pg 2).

Participate in caring for parents
- 19% non-Hispanic Whites
- 28% African Americans
- 34% Hispanic Americans
- 42% Asian Americans

Feel they should do more
- 72% Asian Americans
- 65% Hispanic Americans
- 54% African Americans
- 44% non-Hispanic Whites

Do not expect care from their children
- 72% non-Hispanic Whites
- 68% African Americans
- 60% Hispanic Americans
- 49% Asian Americans

Source: AARP Survey

    (a) Name three things wrong with these figures.

    (b) Sketch an improved version of these figures.