

## Chapter 2: Homework Problems

Take the data from table 2.13, page 39. Answer the following questions on sheets of paper that can be handed in for credit. Be sure your answers are legible and clearly labeled.

- Construct back-to-back leaf-and-stem plots of serum cholesterol values before and after the diet, and leaf-and-stem and box plots in the change in serum cholesterol values. On each of the plots, indicate the median (M or  $Q_2$ ),  $Q_1$ ,  $Q_3$ , the maximum ( $Q_4$ ), and the minimum ( $Q_0$ ) values. On the box plot also indicate the mean.
- Calculate the mean and standard deviation for serum cholesterol before the diet.
- What is the median serum cholesterol before the diet?
  - Is  $n$  for serum cholesterol odd or even?
  - Rank the serum cholesterol values from the lowest to highest.
  - Compute the median serum cholesterol.
- What is the range for serum cholesterol before the diet?
  - What is the definition of R?
  - Calculate R.
- What is the mode for serum cholesterol before the diet?
- What is the coefficient of variation for serum cholesterol before the diet?
  - What is the definition of CV?
  - Calculate the CV.
- Many researchers believe that the distribution of serum cholesterol of Americans is right-skewed. If this sample is a random sample representative of typical Americans, we expect that these data will be right-skewed. Draw a diagram of a right-skewed sample of data.
- Because serum cholesterol is generally right-skewed, researchers compute the geometric mean instead of the arithmetic mean or regular mean.
  - Calculate the geometric mean of serum cholesterol before the diet using the  $\log_{10}$  transformation, where the  
mean of  $\log_{10}(\text{serum cholesterol}) = 1/n \sum_{i=1}^n \log_{10}X_i$   
$$= 1/24 \sum_{i=1}^n \log_{10}(\text{serum cholesterol}_i)$$
  - Is the geometric mean larger or smaller than the arithmetic mean?
  - Calculate the geometric standard deviation of serum cholesterol before the diet using the  $\log_{10}$  transformation. Is the geometric standard deviation larger or smaller than the standard deviation obtained from non-transformed (i.e., raw) data.
- For visually depicting skewed data, we will group the serum cholesterol values into 5 groups of 25 units each:  $I_1 = 125.5-150.5$ ;  $I_2 = 150.5-175.5$ ;  $I_3 = 175.5-200.5$ ;  $I_4 = 200.5-225.5$ ; and  $I_5 = 225.5-250.5$ .
  - Using absolute or relative frequencies, plot the histogram for these groups.

- b) Does the shape of this plot suggest that the data for serum cholesterol before the diet are right-skewed?
  - c) Plot the histogram using cumulative frequencies.
10. Using the grouped data, calculate the grouped mean and grouped standard deviation. How do these values compare with those obtained from non-transformed and transformed values?