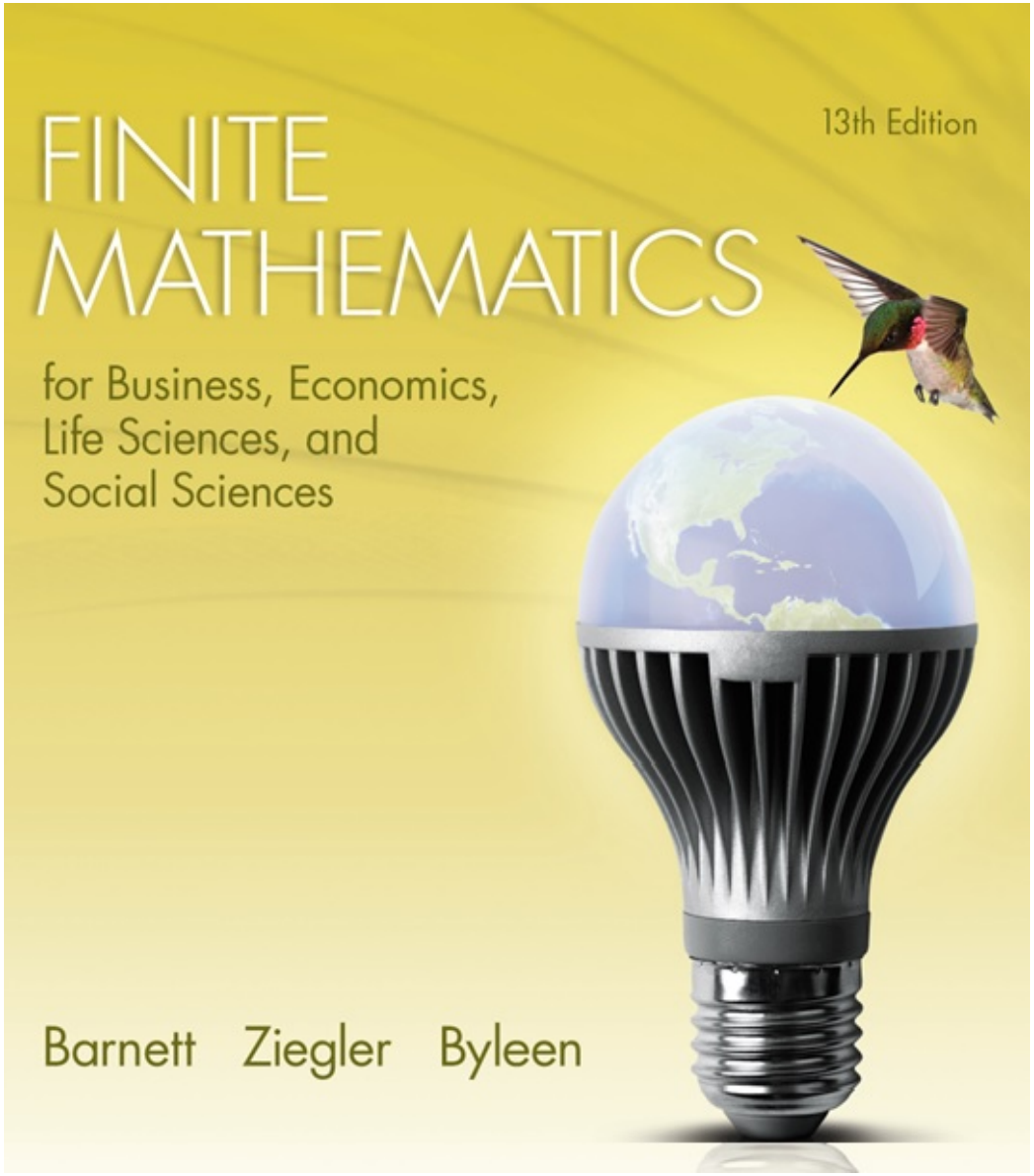

Chapter 11

Data Descriptions and Probability Distributions

Section R

Review

Chapter 11 Review
Important Terms, Symbols, Concepts

- Sect. 11.1 Graphing Data
  - **Bar graphs, broken-line graphs**, and **pie graphs** are used to present visual interpretations or comparisons of data.
  - Large sets of quantitative data can be organized in a **frequency table**, generally constructed by choosing 5 to 20 **class intervals** of equal length to cover the **data range**.
  - The number of measurements that fall in a given class is called the **class frequency**.

Chapter 11 Review

- Section 11.1 Graphing Data (continued)
  - The set of all frequencies associated with their respective classes is called a **frequency distribution**.
  - The **relative frequency** of a class is its frequency divided by the total number of items in the data set.
  - A **histogram** is a vertical bar graph used to represent a frequency distribution.
  - A **frequency polygon** is a broken-line graph obtained by joining successive midpoints of the tops of the bars in a histogram.

Chapter 11 Review

- Section 11.1 Graphing Data (continued)
  - A **cumulative frequency polygon**, or **ogive**, is obtained by plotting the cumulative frequency over the upper boundary of the corresponding class.
- Section 11.2 Measures of Central Tendency
  - The **mean** of a set of quantitative data is the sum of all the measurements in the set divided by the total number of measurements in the set.
  - The **mean for data grouped into classes** is a weighted average of the midpoints of the class intervals.

Chapter 11 Review

- Section 11.2 Measures of Central Tendency (continued)
  - When a data set has *n* measurements and these measurements are arranged in ascending or descending order, the **median** is the middle measurement when *n* is odd, and the mean of the two middle measurements when *n* is even.
  - The **median for grouped data** with no classes of frequency 0 is the number such that the histogram has the same area to the left of the median as to the right of the median.
  - The **mod**e is the most frequently occurring measurement in a data set.

Chapter 11 Review

- 11.3 Measures of Dispersion
  - The **variance** and the **standard deviation** of a set of grouped or ungrouped measurements indicate how the data is dispersed relative to the mean.
- 11.4 Bernoulli Trials and Binomial Distributions
  - A sequence of experiments is called a sequence of **Bernoulli trials**, or a **binomial experiment**, if
    1. Only two outcomes are possible on each trial
    2. The probability of success p is the same for each trial.
    3. All trials are independent.

Chapter 11 Review

- 11.4 Bernoulli Trials and Binomial Distributions (continued)
  - If the random variable x_n represents the number of successes in n Bernoulli trials, then the probability distribution of x_n is the binomial distribution given by

$$p(x_n = x) = C_{n,x} p^x q^{n-x}$$

  - The mean and standard deviation of a binomial distribution are given by the formulas

$$\mu = np \quad \text{and} \quad \sigma = \sqrt{npq}$$

Chapter 11 Review

- 11.5 Normal Distributions
  - Normal curves are bell-shaped continuous curves that approximate the relative frequency distributions of many different types of measurements.
  - The probability that a normally distributed measurement lies between a and b, denoted $p(a \leq x \leq b)$ is equal to the area under the normal curve from a to b.
  - To approximate a binomial distribution that is associated with a sequence of n Bernoulli trials, each having probability of success p, use a normal distribution with

$$\mu = np \quad \text{and} \quad \sigma = \sqrt{npq}$$