Chapter 5: Exploring Data: Distributions

Section 5.2 Interpreting Histograms

James Baglama
Department of Mathematics
University of Rhode Island

- Examining a Distribution
  - Overall Pattern *What does the histogram graph look like?*
    - Shape –
      - Single peak (either symmetric or skewed distribution)
        - » Symmetric – The right and left sides are mirror images.
        - » Skewed to the right – The right side extends much farther out (like a slide).
        - » Skewed to the left – the left side extends much farther out.
      - Irregular distribution of data may appear clustered and may not show a single peak *(due to more than one individual being graphed)*.
    - Center – Estimated center or midpoint of the data.
    - Spread – The range of data outcomes *(minimum to maximum)*.
  - Deviation *Are there any striking differences from the pattern?*
    - Outlier – An individual value that clearly falls outside the overall pattern; possibly an error or some logical explanation.

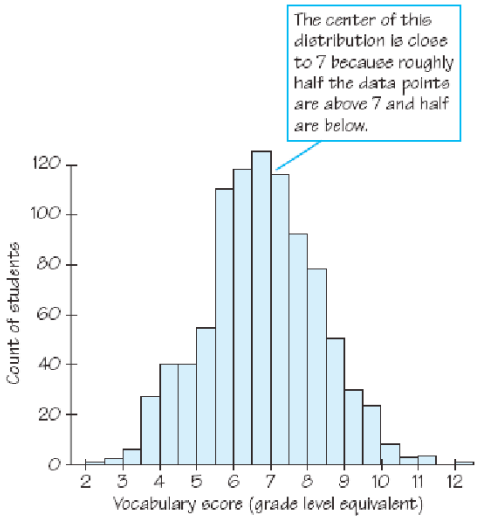
2

Examples of Distribution Patterns and Deviations

Regular Single-Peak Distributions

Histogram of Iowa Test of Basic Skills vocabulary scores for 947 seventh-grade students

Single Peak
Symmetric →



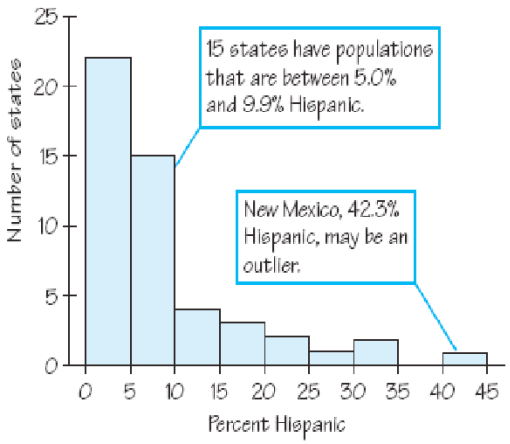
3

Examples of Distribution Patterns and Deviations

Regular Single-Peak Distributions

Histogram of the percent of Hispanics among the adult residents of the states

Single Peak Skewed to Right with Outlier →



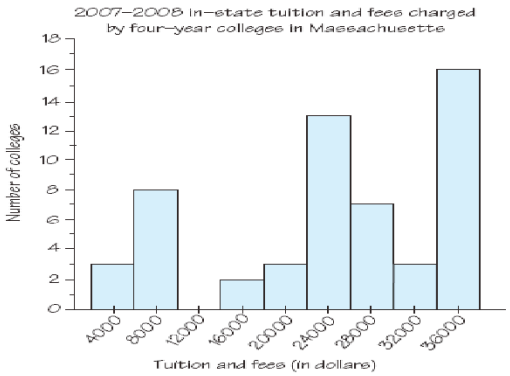
4

Examples of Distribution Patterns and Deviations

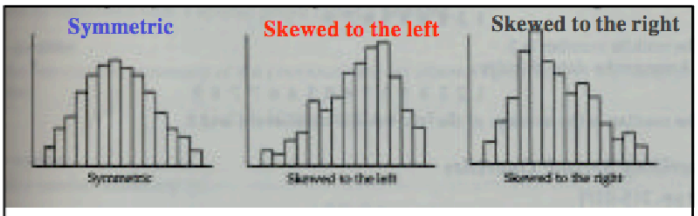
- Irregular Clustered Distributions

Histogram of the tuition and fees charged by four-year colleges in Massachusetts

Two separate distributions, graphing two individuals (state and private schools)→



5

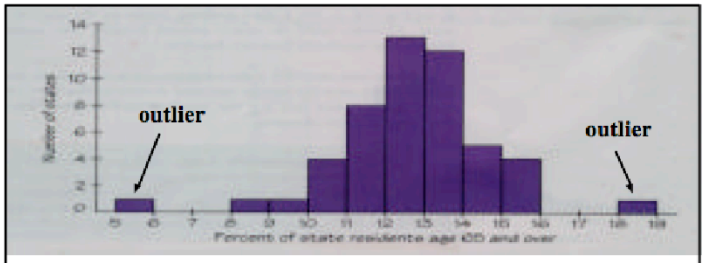


A distribution is **symmetric** if the right and left sides of the histogram are approximately mirror images of each other.

A distribution is **skewed to the left** if the left side of the histogram extends much farther out than the right side.

A distribution is **skewed to the right** if the right side of the histogram extends much farther out than the left side.

6



Describing a Distribution:

- ***Shape:*** the distribution has a single peak. It is roughly symmetric.
- ***Center:*** the midpoint of the distribution is close to the single peak at about 13%.
- ***Spread:*** The spread is about 10% to 16% if we ignore the four most extreme observations.

An outlier is an individual value that falls outside the overall pattern.

7