Chapter Four - Part I
Describing Data: Measures of Dispersion

GOALS
When you have completed this lecture, you will be able to:

ONE
Compute and interpret the range, the variance and the standard deviation.

TWO
Explain the characteristics, uses, advantages, and disadvantages of each measure of dispersion.

THREE
Understand Chebyshev’s Theorem and the Empirical Rule, as they relate to a set of observations.

---

Range

- The range is the difference between the highest and lowest values in a set of data.
  
  \[ \text{RANGE} = \text{Highest Value} - \text{Lowest Value} \]

- EXAMPLE: A sample of five accounting graduates revealed the following starting salaries: $42,000, $38,000, $41,000, $33,000, $44,000.

  \[ \text{RANGE} = \]

---

Population Variance

- The population variance is the average of the squared deviations from the population mean.
  
  \[ \sigma^2 = \frac{\sum (X - \mu)^2}{N} \]
EXAMPLE

The ages of the Dunn family are 2, 18, 34, and 42 years. What is the population variance?

\[ \mu = \frac{\Sigma X}{N} = \frac{96}{4} = 24 \]

\[ \sigma^2 = \frac{\Sigma (X - \mu)^2}{N} = \]

Interpretation: The average squared deviation of the ages from the mean is _________ years² (Huh?)

---

The Population Standard Deviation

- The population standard deviation (\( \sigma \)) is the square root of the population variance.
- For the previous example, the population standard deviation is

\[ \sigma = \sqrt{\sigma^2} = \]

- Interpretation: The “average” deviation of observations from the mean is _________ years.

---

Sample Variance

- The sample variance estimates the population variance.

Conceptual Formula:

\[ s^2 = \frac{\Sigma (X - \bar{X})^2}{n - 1} \]

Formula for Computations:

\[ s^2 = \frac{\Sigma X^2 - (\Sigma X)^2}{n} = \frac{\Sigma X^2 - (\Sigma X)^2}{n - 1} \]

EXAMPLE

- A sample of five hourly wages for various jobs on campus is: $7, $5, $11, $8, $6. Find the variance.

\[ \bar{X} = \frac{7 + 5 + 11 + 8 + 6}{5} = 7.40 \]

- Interpretation: The average squared deviation of observations from the mean is _________ dollars². (Again, huh?)
Sample Standard Deviation

- The sample standard deviation is the square root of the sample variance.
- In the previous example, the sample standard deviation is

\[ s = \sqrt{s^2} = \]

- Interpretation: The “average” deviation of observations from the mean is ________ dollars.

Advantages and Disadvantages of the Measures of Dispersion

- Range:
  - Advantage: Simple to compute.
  - Disadvantage: Not as informative
- Variance:
  - Advantages: Uses all of the sample or population values. Has “good” mathematical properties.
  - Disadvantages: Has squared units for interpretation. Is sensitive to extremely large or small values.

Advantages and Disadvantages of the Measures of Dispersion

- Standard Deviation:
  - Advantages: Uses all of the sample or population values. Has the same units as the original data. Measures the distance of values from the mean
  - Disadvantages: Is sensitive to extremely large or small values.

Interpretation and Uses of the Standard Deviation

- Chebyshev’s Theorem: For any set of observations, the minimum proportion of the values that lie within \( k \) standard deviations of the mean is at least \( 1 - 1/k^2 \), where \( k^2 \) is any constant greater than 1.
EXAMPLE

Recall the example regarding the number of study hours for ST204 students. The mean of the data set was approximately 9.5 hours and the standard deviation was approximately 5.0 hours. At least what percentage of ST204 students study between 3.5 and 15.5 hours?

EXAMPLE

First we need to determine k:

Then we use Chebyshev’s Theorem to get

\[ 1 - \frac{1}{k^2} = \]

Interpretation: At least ____% of the study hours will be between 3.5 hours and 15.5 hours.

Interpretation and Uses of the Standard Deviation

Empirical Rule: For any symmetrical, bell-shaped distribution,

- approximately 68% of the values will lie within 1σ of the mean (μ);
- approximately 95% of the values will lie within 2σ of the mean (μ);
- approximately 99.7% of the values will lie within 3σ of the mean (μ).
EXAMPLE

Consider a factory that produces 2.5 inch bolts. The distribution of bolt lengths is shown on the next slide. According to the Empirical Rule, approximately what percentage of the bolts should measure between 2.49 and 2.51 inches in length?

DISTRIBUTION OF BOLT LENGTHS FOR PROCESS 2.

The End of Chapter 4
Part I

Second part of
Textbook Assignment 2:
Page 103 - #8,
Page 106 - #14 (see pg 105 for an example),
Page 112 - #24, 26.