Chapter 5: Measures of Variability

"An approximate answer to the right question is worth a good deal more than the exact answer to an approximate question."
-- John Tukey

Learning Objectives

Upon completion of this chapter, students should know

- The function of variability measures.
- How to define and calculate the range, average mean deviation, variance, and standard deviation.
- Why a correction factor is necessary when sample variance is used to estimate population variance.
- How to interpret measures of variability.

Key Terms

Measures of variability indicate how much scores in a distribution vary either from the mean or the full extent of the distribution. It is the spread of all the scores in a distribution. Four measures of variability are discussed in this chapter: the range, the average mean deviation, the variance, and the standard deviation. These measures of variability can reveal the consistency or similarity of the scores in a distribution and the extent the mean truly represents all of the scores in the distribution.

The **range** is a measure of the full extent of how far the scores are spread out in a distribution from the highest to the lowest. It is the easiest of the measures of variability to compute, but it is seldom used because of its instability. One extreme score can drastically alter the range.
Range = high score – low score.

Average mean deviation (*AMD*) is the average deviation of each deviation score from the mean of the distribution. To compute the average mean deviation, the mean is subtracted from each score to arrive at the **deviation scores**. The deviation scores are summed and then divided by the total number of scores. Since about half of the deviation scores are positive for the scores larger than the mean and the other half are negative for scores smaller than the mean, the average mean deviation is always equals zero. Thus, because the result is always zero, it cannot be used for any meaningful comparisons between different distributions.

$$\text{Average mean deviation } (AMD) = \frac{\sum (X - \bar{X})}{n} = \frac{\sum x}{n}$$

Variance is represented by either s^2 (population) or S^2 (sample). Variance is computed from deviation scores that are squared to remove the positive and negative signs. The mean of the squared deviation scores is called the variance.

Population variance is the actual computed variance of the population. Sample variance is an estimate of the population variance using only the values in the sample. Because samples rarely contain all the extreme

scores in a population, their variances are generally smaller than those of the population. To be a good estimate of the variance of the population, a correction factor of $n - 1$ is used to increase the variance of a sample.

$$S^2 = \frac{\sum X^2 - \frac{\sum (X)^2}{n}}{n-1}$$

Standard deviation is represent by either s (population) or S (sample). The standard deviation is the square root of the variance. It represents the average amount each score in the distribution deviates from the mean. It is expressed in the same units as the original scores and indicates the consistency or similarity of the scores in a distribution and the extent the mean truly represents all of the scores in the distribution. $S = \sqrt{S^2}$

Lecture and Demonstration Ideas

Students should know that measures of central tendency provide useful information about the characteristics of the distribution, but the description is incomplete without describing its variability.

In addition to providing a more complete picture of a distribution, a conceptual understanding of variance and standard deviation are especially important to understanding normal distributions as well as a central concept in most of the statistical tests in this course.

Conceptually Evaluating Distribution Differences. Before introducing the topic, ask students to evaluate and describe the differences between the two distributions of data (see Transparency 5-2) with the same mean, median, and range. Most likely, their evaluations will involve subtracting the mean from the other scores in the distributions to arrive at an average variability measure. Your demonstration (Transparencies 5-1 to 5-7) and the results that correspond to their evaluations will reinforce their use of this reasoning. This will help students begin to conceptually relate central tendency measures and variability to normal curves and the transformation rules used in z-scores.

You may want to take this opportunity use the Sample B data to discuss the reemphasize the importance of visually examining data before making judgments about central tendency measures. In this case, was the mean the best measure? In a bimodal distribution, is the variability described relating to the mean a fair representation of the distribution?

Instructional Video. *Against All Odds: Inside Statistics*. Program Three, “Numerical Descriptions of Distributions” has an example of median, mean, and quartiles and salary. You may want to show the first two segments on central tendency and skewed distributions now and show the last two segments when you discuss measures of dispersion. The videos are produced by the Consortium for Mathematics and Its Applications and Chedd-Angier (1989) and available through Annenberg/CPB.

Active-Learning Activities

Revisit the Psychologists’ Salary Data. Revisit the 1999 salary data of psychologists (APA, 2000) and discuss standard deviation (Transparency 4-5).

Television Violence. Have students compute the variance and standard deviation of the rate of violence in television outlets (Center for Media Public Affairs, 1999). The full solution is shown below.

Television Violence Study – Rate of Violence Per Episode			
Lifetime	1	Basic Cable	10
NBC	2	Premium Cable	10
PAX	2	S/T All Cable	10
ABC	3	Sci-Fi	11
MTV	3	USA	14
Family	3	CBS	16
FOX	4	HBO	31
Showtime	4	Syndicated	35
WB	7	TNT	73
UPN	8		
Mean = 13.00 scenes per episode Variance = 297.667 Standard deviation = 17.253 scenes per episode		$\sqrt{\frac{8569-\frac{247^2}{19}}{18}} = \sqrt{\frac{8569-3211}{18}} = \sqrt{297.667} = 17.253$	

Height Demonstration. Ask students to compute the range, variance, and standard deviation using the height data collected during the height distribution demonstration of central tendency in your class or the class data collected from one of my classes shown in Handout 5-A. The full solution this dataset is shown on Transparency 5-10.

Handout 5-A.

Height of Men (in inches)

65	68	70	70	72
66.5	69	70	71	72
67	69	70	71	72
68	69	70	71	73
68	69	70	71	73
68	69	70	72.5	

Compute the mean, range, variance, and standard deviation for the distribution of height listed above.

Transparency 5-1.

Measures of Dispersion

- ✓ Range
- ✓ Average Mean Deviation
- ✓ Variance
- ✓ Standard Deviation

What does it indicate?

The similarity of the scores in a distribution.

The extent the mean represents the scores in the distribution.

Transparency 5-2.

Two Samples

Class A	Class B
1	1
5	1
5	1
5	5
5	9
5	9
9	9

Range A = 9 – 1 = 8

Mean A = $\frac{35}{7} = 5$

Median A = 5

Range B = 9 – 1 = 8

Mean B = $\frac{35}{7} = 5$

Median B = 5

Transparency 5-3.

Sample A – Average Mean Deviation

X	$(X - \bar{X})$	**x**
1	1 – 5	-4
5	5 – 5	0
5	5 – 5	0
5	5 – 5	0
5	5 – 5	0
5	5 – 5	0
9	9 - 5	4
		$\sum x = 0$

$n = 7$

$$\frac{0}{7} = 0$$

AMD = 0

Transparency 5-4.

Sample A

Get to the root of the problem – square X.

X	$(X-\bar{X})$	x	x^2
1	1 – 5	-4	16
5	5 – 5	0	0
5	5 – 5	0	0
5	5 – 5	0	0
5	5 – 5	0	0
5	5 – 5	0	0
9	9 - 5	4	16

$$\bar{X} = \frac{\sum X}{n} = \frac{35}{7} = 5 \qquad \sum x^2 = 32$$

$$\frac{\sum x^2}{n} = \frac{32}{7} = 4.571$$

Allow for correction factor $n-1$

$$\frac{32}{n-1} = \frac{32}{6} = 5.333$$

Transparency 5-5

Variance and Standard Deviation

(Computation)

Sample A

$$S^2 = \frac{\sum X^2 - \frac{\sum (X)^2}{n}}{n-1} = \frac{207 - \frac{(35)^2}{7}}{7-1} = \frac{207-175}{6} = 5.333$$

$$S = \sqrt{S^2} = \sqrt{5.333} = \mathbf{2.309}$$

Transparency 5-6.

Variance and Standard Deviation

(Computation)

Sample B

$$S^2 = \frac{\sum X^2 - \frac{\sum (X)^2}{n}}{n-1} = \frac{271 - \frac{(35)^2}{7}}{7-1} = \frac{271-175}{6} = 16$$

$$S = \sqrt{S^2} = \sqrt{16.00} = 4.00 = \mathbf{2.309}$$

Transparency 5-7

Average Distance of Scores From the Mean

Sample A	Sample B
1	1
5	1
5	1
5	5
5	9
5	9
9	9

Mean A = $\frac{35}{7} = 5$

Mean B = $\frac{35}{7} = 5$

$$S = \sqrt{S^2} = \sqrt{5.333}$$

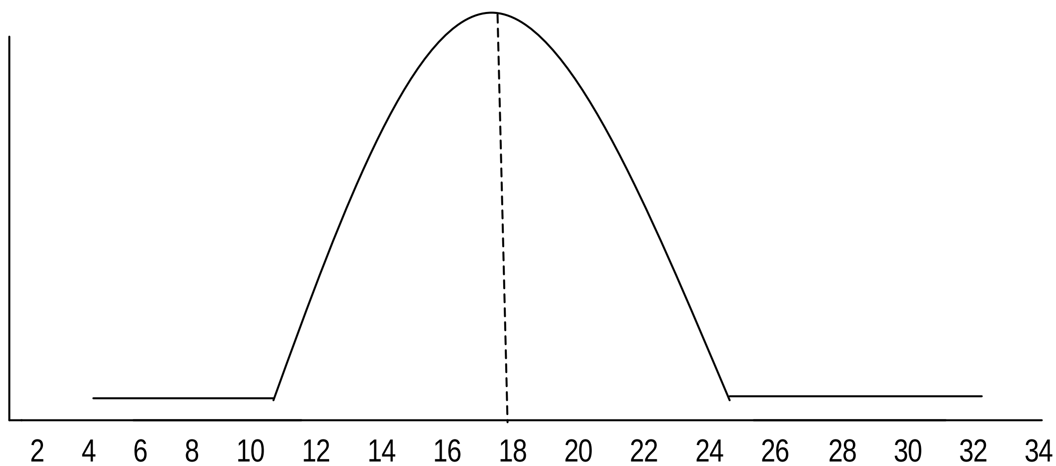
= 2.309

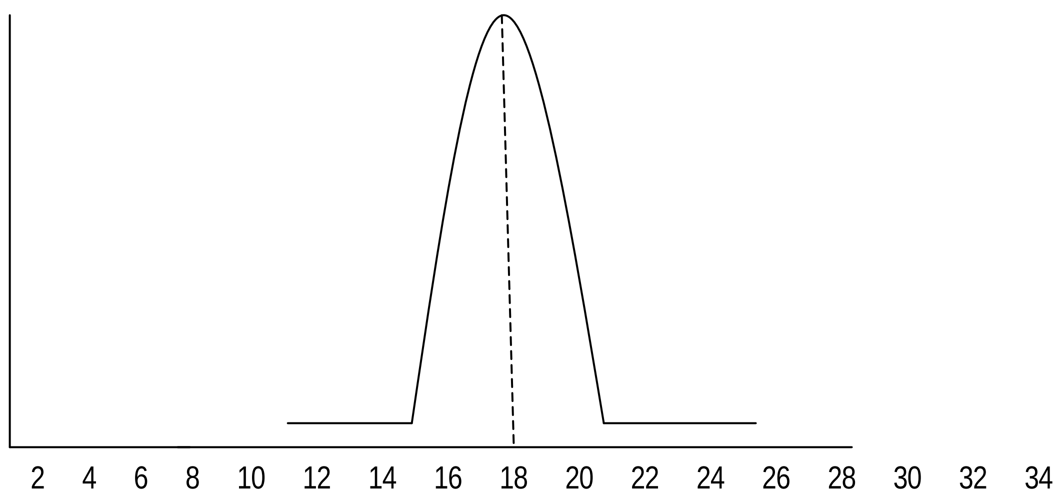
$$S = \sqrt{S^2} = \sqrt{16.00}$$

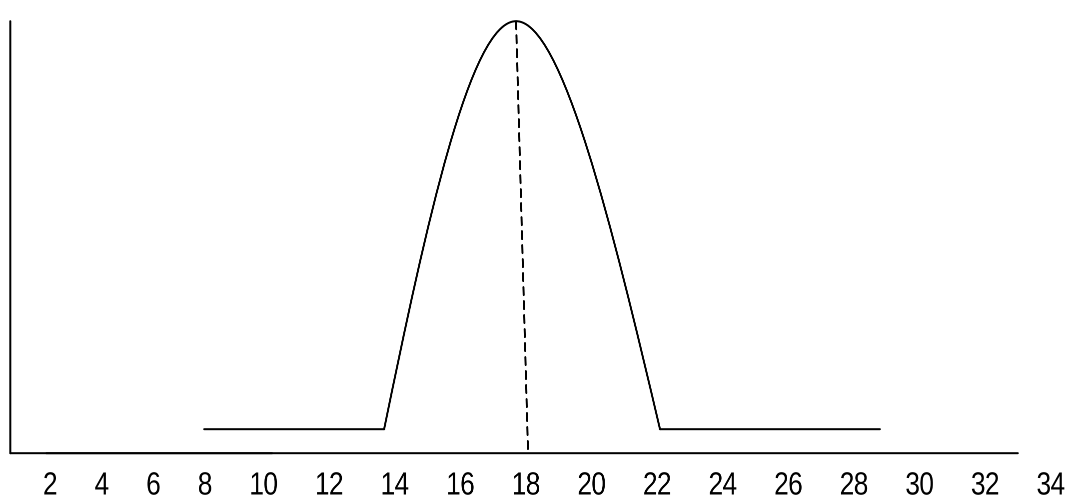
= 4.00

Transparency 5.8

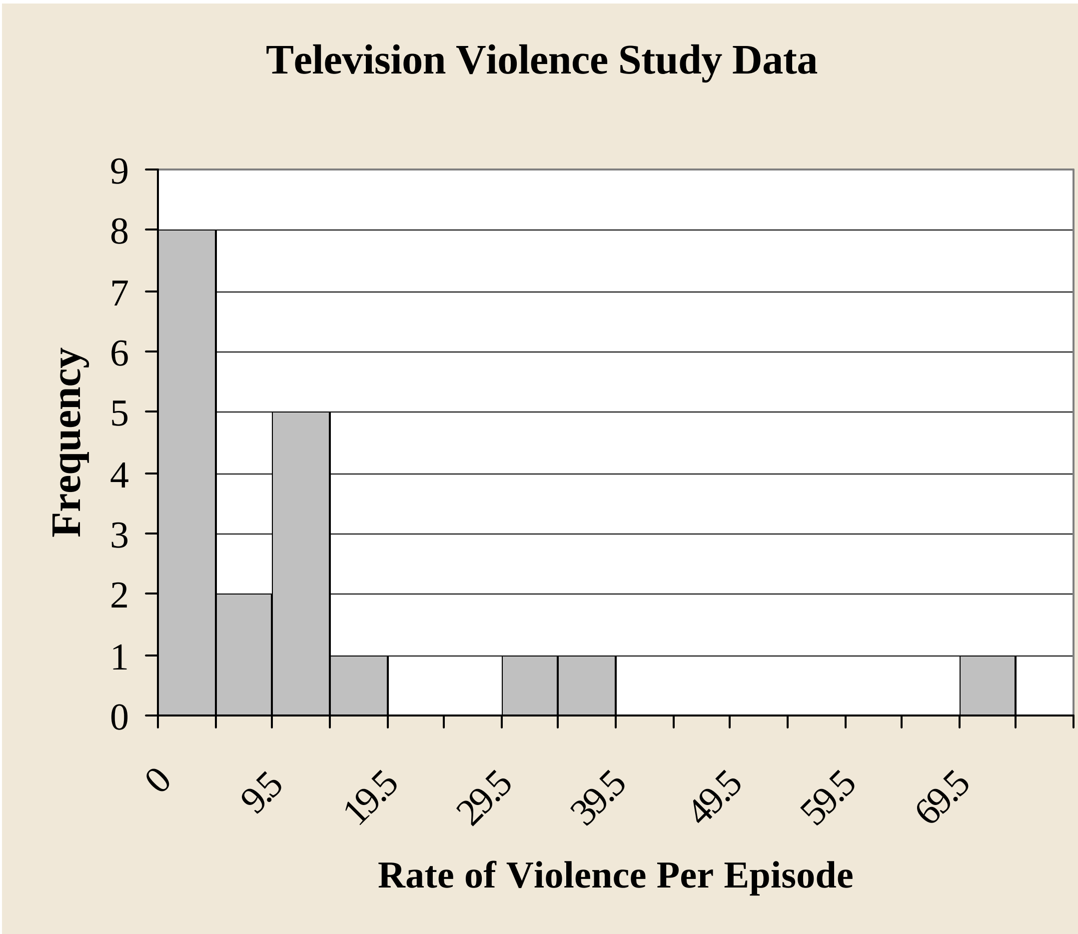
Which distribution has the least variability in scores?



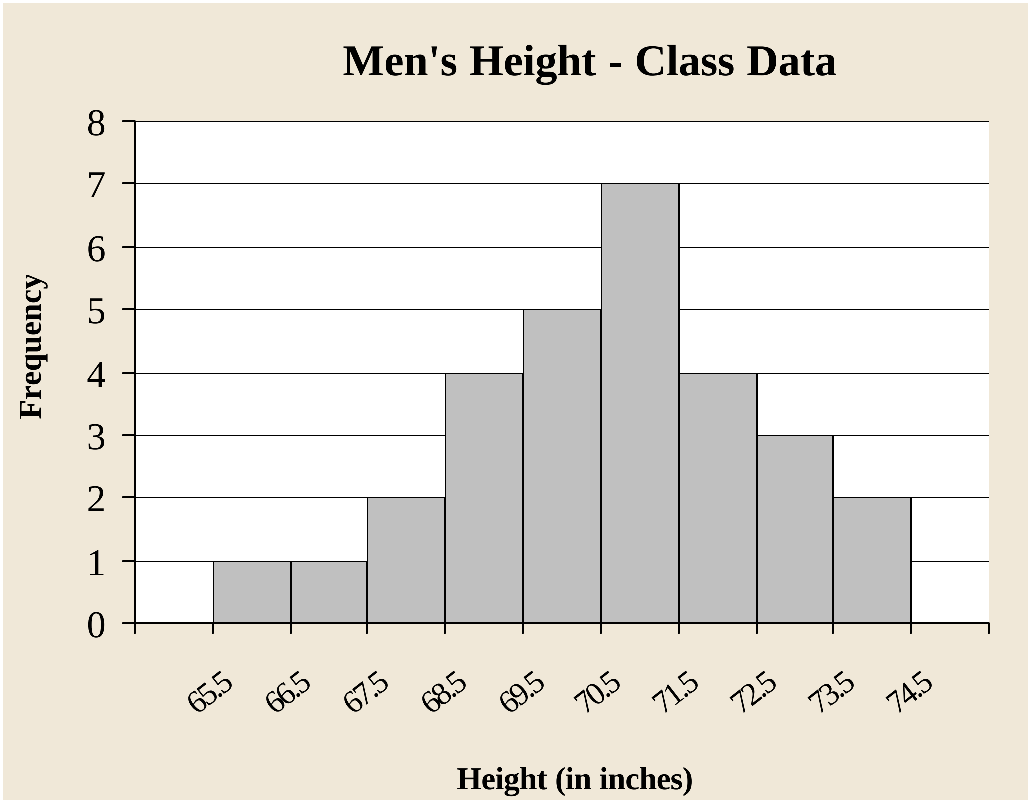




Transparency 5-9.



Transparency 5-10.



Summary data = $\sum X = 2{,}024$, $\sum X^2 = 141{,}367.5$, $n = 29$

$\bar{X} = \dfrac{2024}{29} = 69.793$ inches Range = 73 – 54 = 8 inches

$$S^2 = \frac{141{,}367.5 - \dfrac{2024^2}{29}}{29-1} = \frac{141{,}367.5 - 141{,}261.24}{28} = 3.795$$

$$S = \sqrt{3.795} = 1.948 \text{ inches}$$