04

Chapter

Chapter 4
Measures of Variability

PowerPoint Lecture Slides

Essentials of Statistics for the Behavioral Sciences

Eighth Edition

by Frederick J Gravetter and Larry B. Wallnau

04 Chapter

Learning Outcomes

1. Understand purpose of measuring variability
2. Define range
3. Compute range
4. Understand variance and standard deviation
5. Calculate SS, variance, standard deviation of population
6. Calculate SS, variance, standard deviation of sample

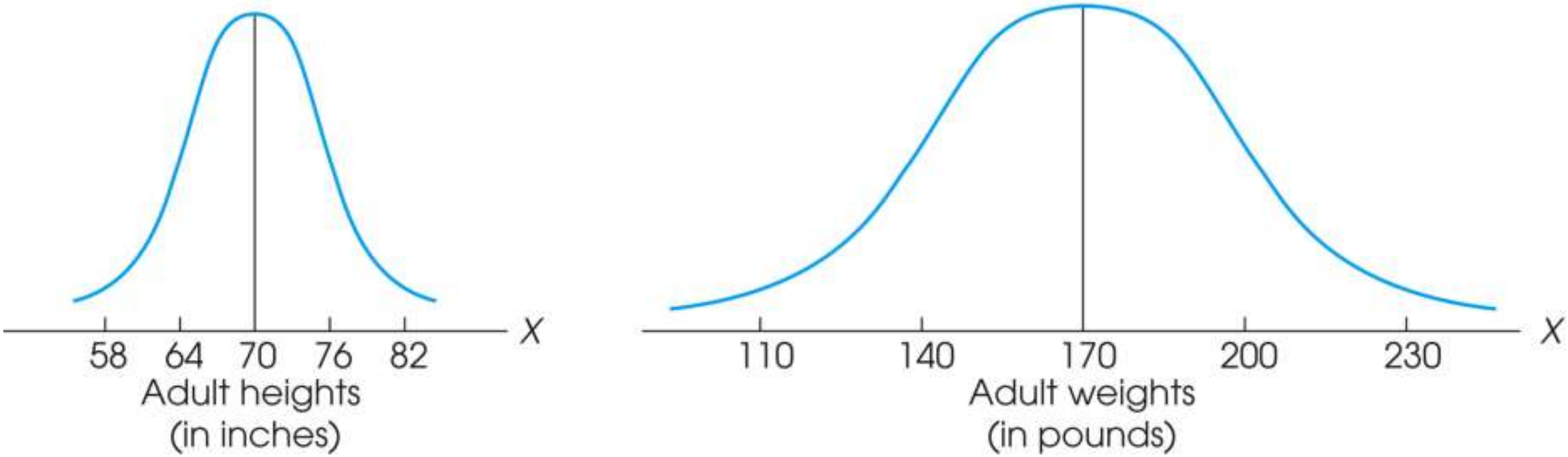
Tools You Will Need

- Summation notation (Chapter 1)
- Central tendency (Chapter 3)
  - Mean
  - Median

4.1 Overview

- Variability can be defined several ways
  - A quantitative *distance* measure based on the differences between scores
  - Describes *distance* of the spread of scores or *distance* of a score from the mean
- Purposes of Measure of Variability
  - Describe the distribution
  - Measure how well an individual score represents the distribution

Figure 4.1 Population Distributions



04 Chapter

Three Measures of Variability

- The Range
- The Variance
- The Standard Deviation

4.2 The Range

- The distance covered by the scores in a distribution
  - From smallest value to highest value
- For continuous data, real limits are used

$$\text{range} = \text{URL for } X_{max} - \mathit{LRL\ for\ } X_{min}$$

- Based on two scores, not all the data
  - An imprecise, unreliable measure of variability

4.3 Standard Deviation and Variance for a Population

- Most common and most important measure of variability is the standard deviation
  - A measure of the standard, or average, distance from the mean
  - Describes whether the scores are clustered closely around the mean or are widely scattered
- Calculation differs for population and samples
- Variance is a necessary ***companion concept*** to standard deviation but ***not*** the same concept

Defining the Standard Deviation

- Step One: Determine the Deviation
- Deviation is distance from the mean

Deviation score $= X - \mu$

- Step Two: Find a "sum of deviations" to use as a basis of finding an "average deviation"
  - Two problems
    - Deviations sum to 0 (because M is balance point)
    - If sum always 0, "Mean Deviation" will always be 0.
  - Need a new strategy!

Defining the Standard Deviation (continued)

- Step Two *Revised*: Remove negative deviations
  - ***First*** square each deviation score
  - Then sum the *Squared* Deviations (SS)
- Step Three: Average the *squared* deviations
  - Mean Squared Deviation is known as "***Variance***"
  - Variability is now measured in squared units

Population variance equals mean (average) squared deviation (distance) of the scores from the population mean

Defining the Standard Deviation (continued)

- Step Four:
  - Goal: to compute a measure of the “standard” (average) distance of the scores from the mean
  - Variance measures the average *squared* distance from the mean—not quite our goal
- Adjust for having squared all the differences by taking the square root of the variance
- $Standard\ Deviation = \sqrt{Variance}$

Figure 4.2

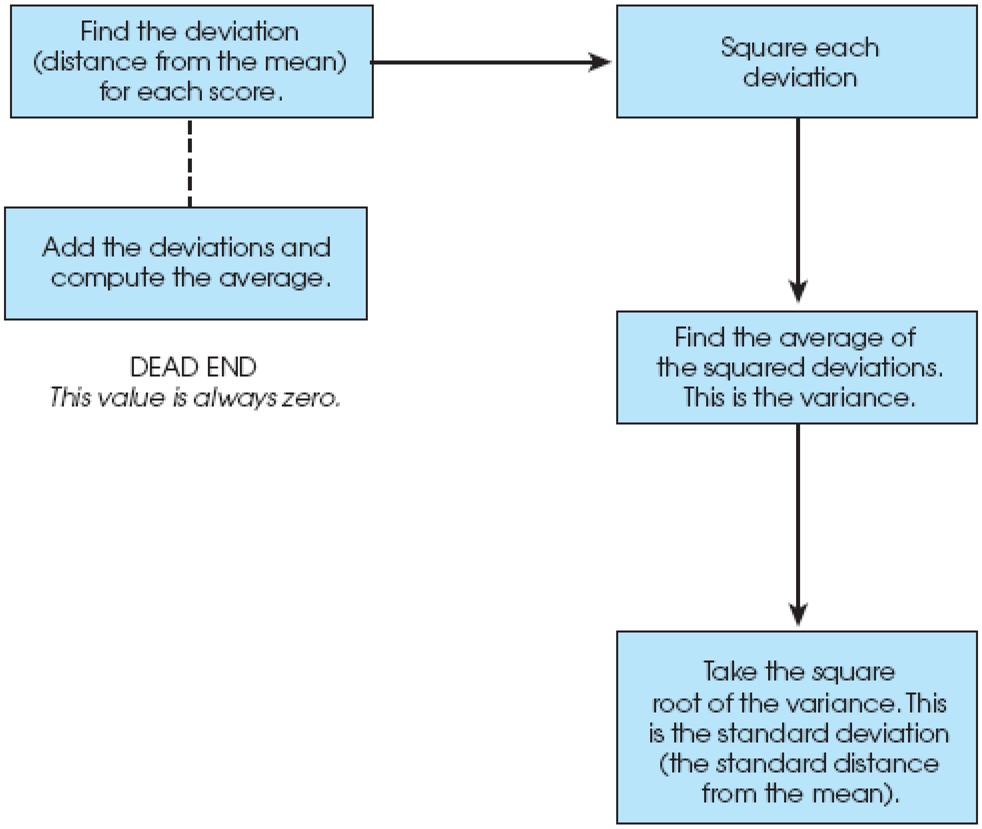
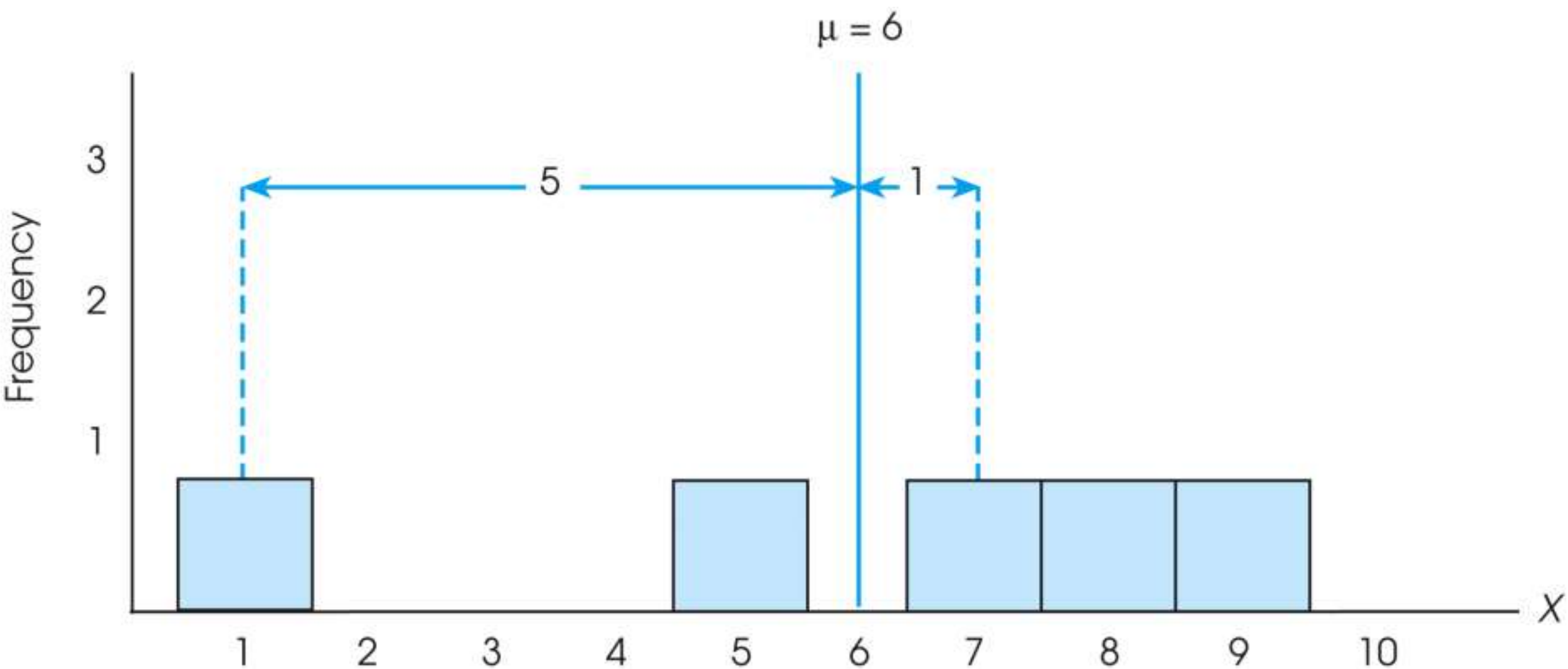

Find the deviation
(distance from the mean)
for each score.
Square each
deviation
Add the deviations and
compute the average.
DEAD END
This value is always zero.
Find the average of
the squared deviations.
This is the variance.
Take the square
root of the variance. This
is the standard deviation
(the standard distance
from the mean).

Figure 4.3



Population Variance Formula

$$\text{Variance} = \frac{\text{sum of squared deviations}}{\text{number of scores}}$$

- SS (sum of squares) is the sum of the squared deviations of scores from the mean
- Two formulas for computing SS

Two formulas for SS

Definitional Formula

- Find each deviation score $(X–\mu)$
- Square each deviation score, $(X–\mu)^2$
- Sum up the squared deviations

$$SS = \sum (X - \mu)^2$$

Computational Formula

- Square each score and sum the squared scores
- Find the sum of scores, square it, divide by N
- Subtract the second part from the first

$$SS = \sum X^2 - \frac{\left(\sum X\right)^2}{N}$$

Caution Required!

When using the computational formula, remember…

$$SS \neq \sum X^2$$

$$SS \neq (\sum X)^2$$

$$\sum X^2 \neq (\sum X)^2$$

Population Variance: Formula and Notation

Formula

$$\textit{variance} = \frac{SS}{N}$$

$$\textit{standard deviation} = \sqrt{\frac{SS}{N}}$$

Notation

- Variance is the average of ***squared*** deviations, so we identify population variance with a lowercase Greek letter sigma squared: σ^2
- Standard deviation is the square root of the variance, so we identify it with a lowercase Greek letter sigma: σ

Learning Check

- Decide if each of the following statements is True or False.

T/F

- The computational & definitional formulas for *SS* sometimes give different results

T/F

- If all the scores in a data set are the same, the Standard Deviation is equal to 1.00

Learning Check - Answer

False

- The computational formula is just an algebraic rearrangement of the definitional formula. Results are identical

False

- When all the scores are the same, they are all equal to the mean. Their deviations = 0, as does their Standard Deviation

Learning Check

- The standard deviation measures …

A
- Sum of squared deviation scores

B
- Standard distance of a score from the mean

C
- Average deviation of a score from the mean

D
- Average squared distance of a score from the mean

Learning Check - Answer

- The standard deviation measures …

A
- Sum of squared deviation scores

B
- Standard distance of a score from the mean

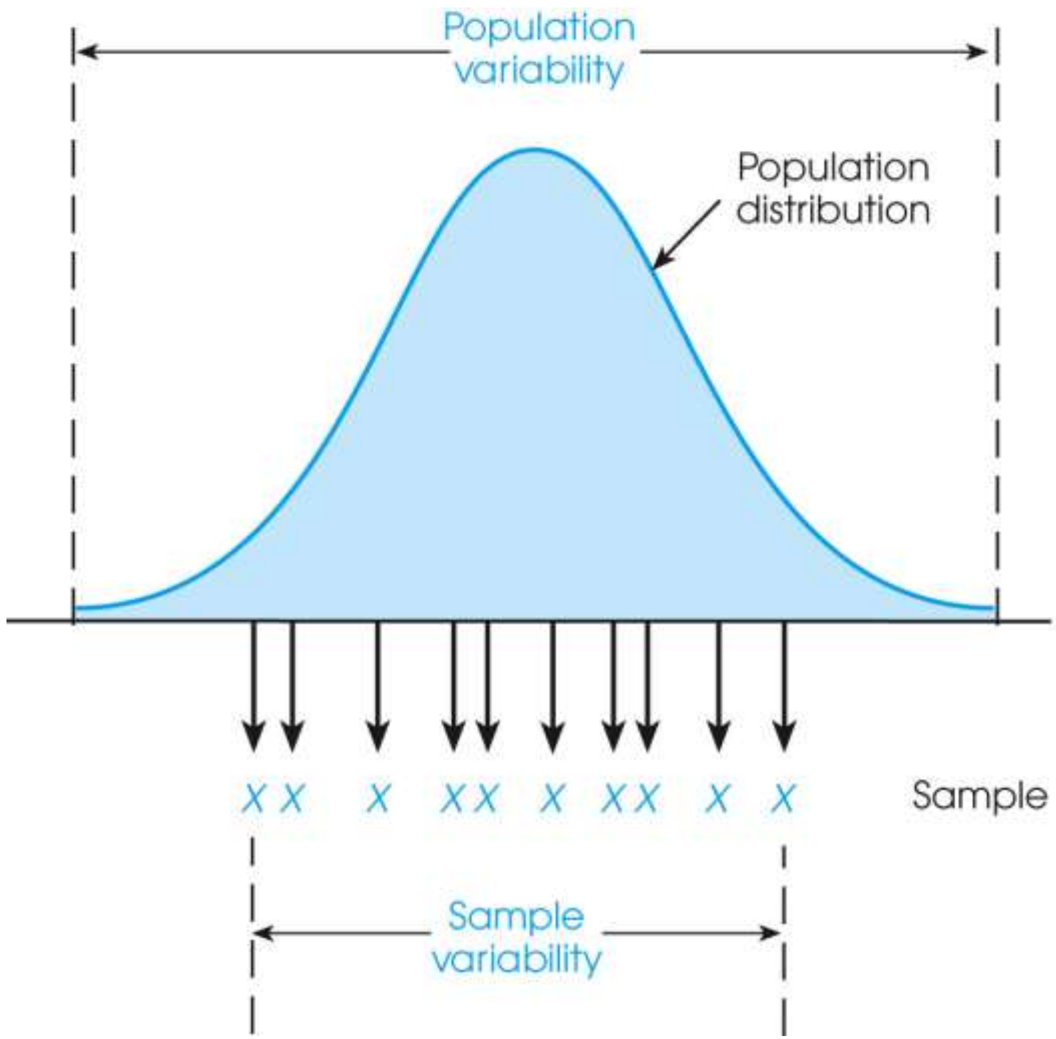
C
- Average deviation of a score from the mean

D
- Average squared distance of a score from the mean

4.4 Standard Deviation and Variance for a Sample

- Goal of inferential statistics:
  - Draw general conclusions about population
  - Based on limited information from a sample
- Samples differ from the population
  - Samples have less variability
  - Computing the Variance and Standard Deviation in the same way as for a population would give a biased estimate of the population values

Figure 4.4 Population of Adult Heights



04 Chapter

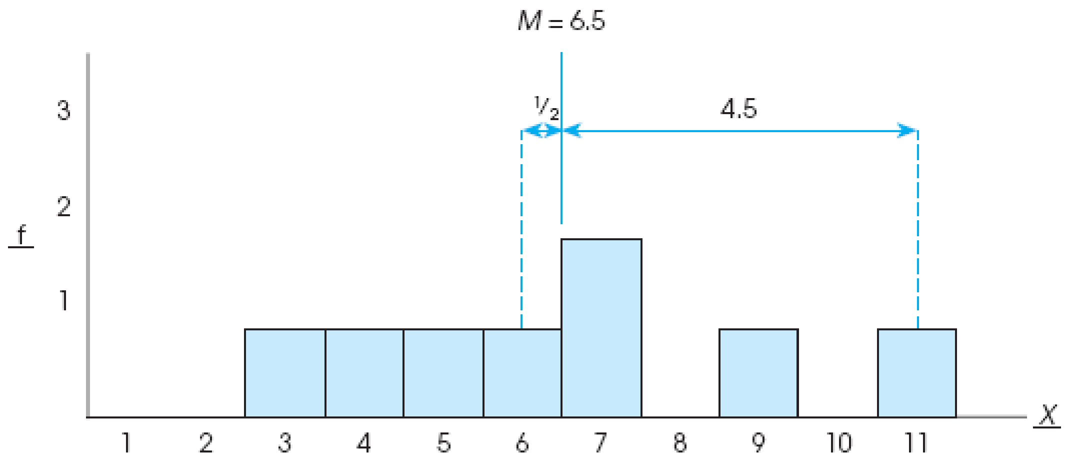
Sample Variance and Standard Deviation

- Sum of Squares (SS) is computed as before
- Formula for Variance has n-1 rather than N in the denominator
- Notation uses *s* instead of σ

$$variance\ of\ sample = s^2 = \frac{SS}{n-1}$$

$$standard\ deviation\ of\ sample = s = \sqrt{\frac{SS}{n-1}}$$

Figure 4.5
Histogram for Sample of n = 8



Degrees of Freedom

- Population variance
  - Mean is known
  - Deviations are computed from a known mean
- Sample variance as estimate of population
  - Population mean is unknown
  - Using sample mean restricts variability
- Degrees of freedom
  - Number of scores in sample that are independent and free to vary
  - Degrees of freedom (df) = n – 1

Learning Check

- A sample of four scores has SS = 24. What is the variance?

A	The variance is 6
B	The variance is 7
C	The variance is 8
D	The variance is 12

Learning Check - Answer

- A sample of four scores has SS = 24.
  What is the variance?

A	The variance is 6
B	The variance is 7
C	**The variance is 8**
D	The variance is 12

Learning Check

- Decide if each of the following statements is True or False.

T/F

- A sample systematically has less variability than a population

T/F

- The standard deviation is the distance from the Mean to the farthest point on the distribution curve

Learning Check - Answer

True

- Extreme scores affect variability, but are less likely to be included in a sample

False

- The standard deviation extends from the mean approximately halfway to the most extreme score

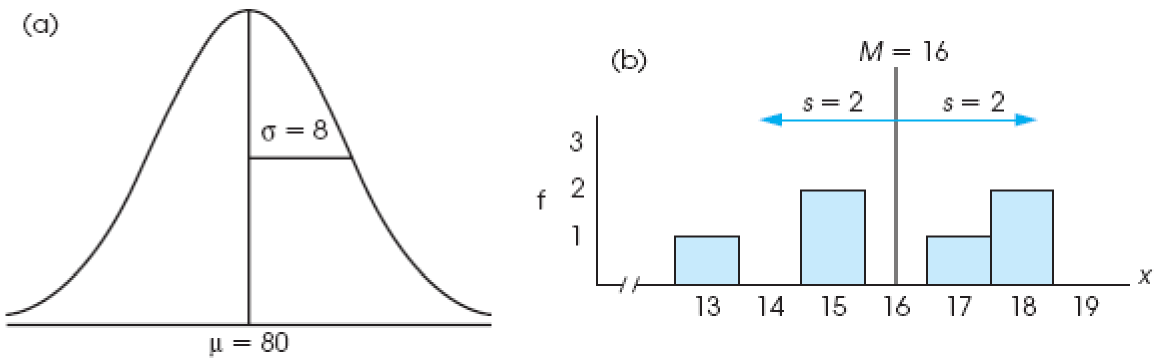
4.5 More About Variance and Standard Deviation

- Mean and standard deviation are particularly useful in clarifying graphs of distributions
- Biased and unbiased statistics
- Means and standard deviations together provide extremely useful descriptive statistics for characterizing distributions

Showing Mean and Standard Deviation in a Graph

- For both populations and samples it is easy to represent mean and standard deviation
  - Vertical line in the “center” denotes location of mean
  - Horizontal line to right, left (or both) denotes the distance of one standard deviation

Figure 4.6 Showing Means and Standard Deviations in Graphs



04 Chapter

Sample Variance as an Unbiased Statistic

- Unbiased estimate of a population parameter
  - Average value of statistic is equal to parameter
  - Average value uses all possible samples of a particular size n
  - Corrected standard deviation formula (dividing by *n*-1) produces an *unbiased estimate* of the population variance
- Biased estimate of a population parameter
  - Systematically overestimates or underestimates the population parameter

Table 4.1 Biased and Unbiased Estimates

			Sample Statistics		
Sample	1st Score	2nd Score	(Unbiased) Mean	Biased Variance	Unbiased Variance
1	0	0	0.00	0.00	0.00
2	0	3	1.50	2.25	4.50
3	0	9	4.50	20.25	40.50
4	3	0	1.50	2.25	4.50
5	3	3	3.00	0.00	0.00
6	3	9	6.00	9.00	18.00
7	9	0	4.50	20.25	40.50
8	9	3	6.00	9.00	18.00
9	9	9	9.00	0.00	0.00
		Totals	**36.00**	**63.00**	**126.00**

Mean of 9 unbiased sample mean estimates: 36/9 = 4 (Actual $\mu = 4$)

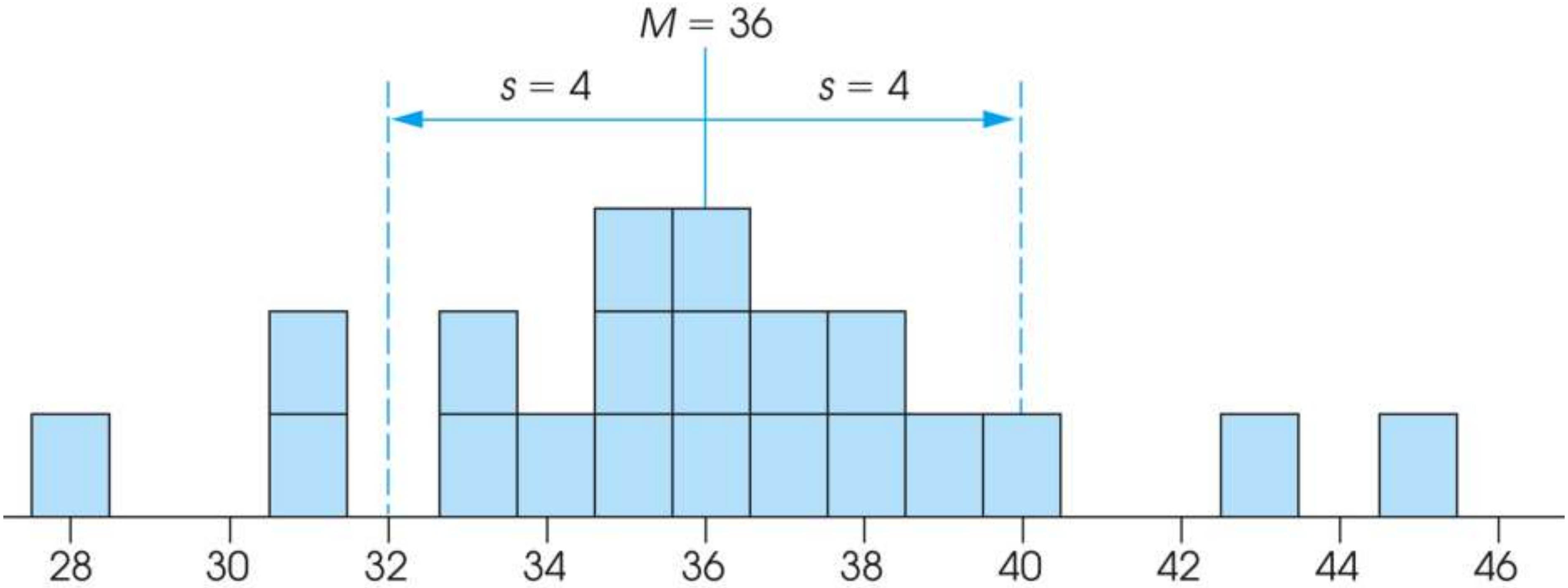
Mean of 9 biased sample variance estimates: 63/9 = 7 (Actual $\sigma^2 = 14$)

Mean of 9 unbiased sample variance estimates: 126/9=14 (Actual $\sigma^2 = 14$)

Standard Deviation and Descriptive Statistics

- A standard deviation describes scores in terms of *distance from the mean*
- Describe an entire distribution with just two numbers (*M* and *s*)
- Reference to both allows reconstruction of the measurement scale from just these two numbers (Figure 4.7)

Figure 4.7 (Sample) $n = 20$, $M = 36$, and $s = 4$



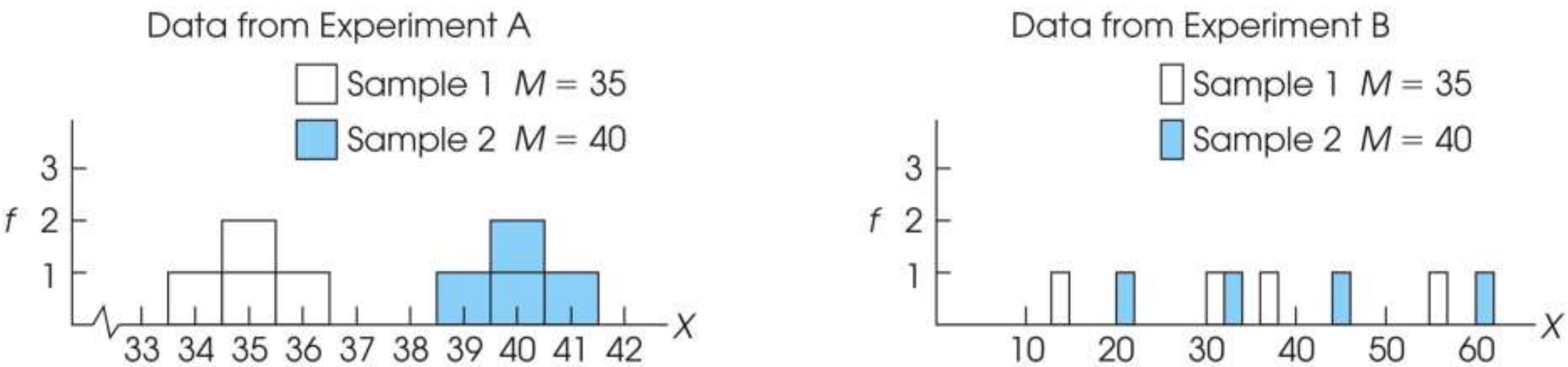
Transformations of Scale

- Adding a constant to each score
  - The Mean is changed
  - The standard deviation is unchanged
- Multiplying each score by a constant
  - The Mean is changed
  - Standard Deviation is also changed
  - The Standard Deviation is multiplied by that constant

Variance and Inferential Statistics

- Goal of inferential statistics is to detect meaningful and significant patterns in research results
- Variability in the data influences how easy it is to see patterns
  - High variability obscures patterns that would be visible in low variability samples
  - Variability is sometimes called error variance

Figure 4.8 Experiments with high and low variability



Learning Check

- A population has $\mu = 6$ and $\sigma = 2$. Each score is multiplied by 10. What is the shape of the resulting distribution?

A
- $\mu = 60$ and $\sigma = 2$

B
- $\mu = 6$ and $\sigma = 20$

C
- $\mu = 60$ and $\sigma = 20$

D
- $\mu = 6$ and $\sigma = 5$

Learning Check - Answer

- A population has $\mu = 6$ and $\sigma = 2$. Each score is multiplied by 10. What is the shape of the resulting distribution?

A
- $\mu = 60$ and $\sigma = 2$

B
- $\mu = 6$ and $\sigma = 20$

C
- **$\mu = 60$ and $\sigma = 20$**

D
- $\mu = 6$ and $\sigma = 5$

Learning Check TF

- Decide if each of the following statements is True or False.

T/F	• A biased statistic has been influenced by researcher error
T/F	• On average, an unbiased sample statistic has the same value as the population parameter

Learning Check - Answer

False

- Bias refers to the systematic effect of using sample data to estimate a population parameter

True

- Each sample’s statistic differs from the population parameter, but the average of all samples will equal the parameter

Figure 4.9 SPSS Summary Table (n = 8) from Example 4.5

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
VAR00001	8	8.00	3.00	11.00	6.5000	2.61861	6.85714
Valid N (listwise)	8						

04
Chapter