MEASURES OF DISPERSION

Math 1001

Quantitative Skills and Reasoning

Columbus State University

THE STANDARD DEVIATION

- To calculate the standard deviation of *n* numbers, it is helpful to use the following procedure:
  1. Determine the mean of the *n* numbers.
  2. For each number, calculate the deviation (difference) between the number and the mean of the numbers.
  3. Calculate the square of each of the deviations and find the sum of these squared deviations.
  4. If the data is a *population*, then divide the sum by *n*. If the data is a *sample*, then divide the sum by *n* – 1.
  5. Find the square root of the quotient in Step 4.

COLUMBUS STATE UNIVERSITY

FIND THE STANDARD DEVIATION

- The following numbers were obtained by sampling a population:

  3, 4, 7, 11, 15

- Find the standard deviation of the sample.

- Step 1: Find the mean of the numbers: $\bar{x} = \frac{3+4+7+11+15}{5} = \frac{40}{5} = 8$

- Step 2: For each number, calculate the deviation between the number and the mean:

x	$x - \bar{x}$
3	3 – 8 = -5
4	4 – 8 = -4
7	7 – 8 = -1
11	11 – 8 = 3
15	15 – 8 = 7

COLUMBUS STATE UNIVERSITY

FIND THE STANDARD DEVIATION

- The following numbers were obtained by sampling a population:

  3, 4, 7, 11, 15

- Find the standard deviation of the sample.

- Step 3: Calculate the square of each of the deviations from Step 2, and find the sum of these squared deviations.

x	$x - \bar{x}$	$(x - \bar{x})^2$
3	3 – 8 = -5	$(-5)^2 = 25$
4	4 – 8 = -4	$(-4)^2 = 16$
7	7 – 8 = -1	$(-1)^2 = 1$
11	11 – 8 = 3	$(3)^2 = 9$
15	15 – 8 = 7	$(7)^2 = 49$
Sum		100

- Step 4: Because we have a sample of n = 5 values, divide the sum 100 by n – 1, which is 4.

$$\frac{100}{4} = 25$$

- Step 5: The standard deviation of the sample is $s = \sqrt{25}$.
- The standard deviation is s =5.0

COLUMBUS STATE UNIVERSITY

USE STANDARD DEVIATIONS

- A consumer group has tested a sample of 8 size-D batteries from each of 3 companies.
- The results of the tests are shown in the following table.
- According to these tests, which company produces batteries for which the values representing hours of constant use have the smallest standard deviation?

Company	Hours of constant use per battery
A	6.2, 6.4, 7.1, 5.9, 8.4, 5.2, 7.5, 9.3
B	6.8, 6.2, 7.2, 5.9, 7.0, 7.4, 7.3, 8.2
C	6.1, 6.6, 7.3, 5.7, 7.1, 7.6, 7.1, 8.5

COLUMBUS STATE UNIVERSITY

USE STANDARD DEVIATIONS

Company	Hours of constant use per battery
A	6.2, 6.4, 7.1, 5.9, 8.4, 5.2, 7.5, 9.3
B	6.8, 6.2, 7.2, 5.9, 7.0, 7.4, 7.3, 8.2
C	6.1, 6.6, 7.3, 5.7, 7.1, 7.6, 7.1, 8.5

- Standard Deviation for A batteries:
  - Mean $= \frac{6.2+6.4+7.1+5.9+8.4+5.2+7.5+9.3}{8} = 7$
  - $s_1 = \sqrt{\frac{(6.2-7)^2+(6.4-7)^2+(7.1-7)^2+(5.9-7)^2+(8.4-7)^2+(5.2-7)^2+(7.5-7)^2+(9.3-7)^2}{8-1}}$

$= \sqrt{\frac{12.96}{7}} = 1.361$ hours

COLUMBUS STATE UNIVERSITY

USE STANDARD DEVIATIONS

Company	Hours of constant use per battery
A	6.1, 6.5, 7.1, 5.9, 8.3, 5.3, 7.5, 9.3
B	6.8, 6.2, 7.2, 5.9, 7.0, 7.4, 7.3, 8.2
C	6.1, 6.6, 7.3, 5.7, 7.1, 7.6, 7.1, 8.5

- Standard Deviation for B batteries:
  - $\text{Mean} = \frac{6.8+6.2+7.2+5.9+7.0+7.4+7.3+8.2}{8} = 7$
  - $s_1 = \sqrt{\frac{(6.8-7)^2+(6.2-7)^2+(7.2-7)^2+(5.9-7)^2+(7.0-7)^2+(7.4-7)^2+(7.3-7)^2+(8.2-7)^2}{7}}$

$= \sqrt{\frac{3.62}{7}} = 0.719$ hours

USE STANDARD DEVIATIONS

Company	Hours of constant use per battery
A	6.2, 6.4, 7.1, 5.9, 8.3, 5.3, 7.5, 9.3
B	6.8, 6.2, 7.2, 5.9, 7.0, 7.4, 7.3, 8.2
C	6.1, 6.6, 7.3, 5.7, 7.1, 7.6, 7.1, 8.5

- Standard Deviation for C batteries:
  - Mean $= \frac{6.1+6.6+7.3+5.7+7.1+7.6+7.1+8.5}{8} = 7$
  - $s_1 = \sqrt{\frac{(6.1-7)^2+(6.6-7)^2+(7.3-7)^2+(5.7-7)^2+(7.1-7)^2+(7.6-7)^2+(7.1-7)^2+(8.5-7)^2}{7}}$

$= \sqrt{\frac{5.38}{7}} = 0.877$ hours

Columbus State University

USE STANDARD DEVIATIONS

- The batteries from B company have the smallest standard deviation.
- According to these results, the B company produces the most consistent batteries with regard to life expectancy under constant use.

COLUMBUS STATE UNIVERSITY

THE VARIANCE

- A statistic known as the *variance* is also used as a measure of dispersion.
- The **variance** for a given set of data is the square of the standard deviation of the data.
- The mathematical notations that are used to denote standard deviations and variances are shown here:
  - σ is the standard deviation of a population
  - σ^2 is the variance of a population
  - s is the standard deviation of a sample
  - s^2 is the variance of a sample

COLUMBUS STATE UNIVERSITY

FIND THE VARIANCE

- Find the variance for the sample 3, 4, 7, 11, 15
- Recall that this is the sample from a previous example.
  - In that previous example, we found $s = \sqrt{25}$.
  - Variance is the square of the standard deviation, thus the variance is $s^2 = \left(\sqrt{25}\right)^2 = 25$

COLUMBUS STATE UNIVERSITY

THE VARIANCE

- Can the variance of a data set be less than the standard deviation of the data set?

- Yes. The variance is less than the standard deviation whenever the standard deviation is less than 1.

COLUMBUS STATE UNIVERSITY

THE VARIANCE

- Although the variance of a set of data is an important measure of dispersion, it has a disadvantage that is not shared by the standard deviation:
  - the variance does not have the same unit of measurement as the original data.
- For instance, if a set of data consists of times measured in hours, then the variance will be measured in hours squared.
- The standard deviation of this data set is the square root of the variance, and as such it is measured in hours, which is a more intuitive unit of measure.

COLUMBUS STATE UNIVERSITY