AFM
Unit 6 Day 3 Notes – Variance and Standard Deviation

Name Key
Date _____

Measures of Spread

- **spread** – describes how similar or varied the set of observed values are for a particular variable (data item). Measures of spread include the range, quartiles and the interquartile range, variance and standard deviation.

We have already talked about 2 measures of spread and using the 1.5IQR rule to determine if we have an outlier:

- **range** – difference between the highest and lowest value in a data set
- **interquartile range** - $Q_3 - Q_1$

We are now going to look at 2 other measures of spread – Variance and Standard Deviation:

- **deviation (or directed distance)** – difference between a data value and the mean of the data set
- **variance** – the sum of the squares of the deviations, divided by the number of values
- **standard deviation** – the square root of the variance

Measuring the variability of a data set allows a more complete description than just stating a measure of central tendency.

Variance

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

($\bar{x}$: Mean; n: # of values in the data)

To do the calculations shown in the formula above:

1. Find the difference between each data point and the mean.
2. Square each of those values.
3. Add them all together.
4. Divide by the number of total data points.

Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

To do the calculations shown in the formula above:

1. Find the variance.
2. Take the square root of the variance

Example 1

14.1, 15.6, 17.9, 21.4, 31.7, 13.2, 15.8 $= 129.7$

a. Find the mean of the above data.

$\frac{129.7}{7} = 18.5$

b. Find the deviation from the mean.

$14.1 - 18.5 = -4.4$	19.36		$31.7 - 18.5 = 13.2$	174.24
$15.6 - 18.5 = -2.9$	8.41		$13.2 - 18.5 = -5.3$	28.09
$17.9 - 18.5 = -0.6$	0.36		$15.8 - 18.5 = -2.7$	7.29
$21.4 - 18.5 = 2.9$	8.41			

c. Find the variance.

$\frac{246.16}{7} = 35.1657$

d. Find the standard deviation.

$\sqrt{35.1657} \approx 5.93$

Example 2

3.5, 6.7, 12.1, 43.5, 51.7, 23.1, 67.8, 31.2

a. Find the range of the above data.

$67.8 - 3.5 = 64.3$

b. Find the IQR of the above data.

↳ $47.6 - 9.4 = 38.2$

c. Use the IQR to determine if there are any outliers in the data set.

$47.6 + 38.2(1.5) = 104.9$

$9.4 - 38.2(1.5) = -47.9$

No Outliers

Example 3

12, 14, 54, 35, 71, 1, 98, 22, 41

a. Find the outliers in the above data set.

Min: 1

Q1: 13

Med: 35

Q3: 62.5

Max: 98

IQR: 49.5

$62.5 + 49.5(1.5) = 136.75$

$13 - 49.5(1.5) = -61.25$

No Outliers

Calculator Notes:

Under, STAT→Calc→1-Var-Stats

Standard Deviation → σ

Comparing Cars

With all of the talk lately about fuel efficient cars, it's important to see car companies live up to their hype. Here are the top five fuel efficient cars. Each car was test driven 5 times on full tanks of gas to see how many miles per gallon they got. The results are as follows

Hyundai Elantra Predicted - 33	*Honda Civic* Predicted - 30	*Mazda 3* Predicted - 28	*Ford Focus* Predicted - 32	*Toyota Corolla* Predicted - 30
31.2	28.8	28.3	33.2	31.2
29.9	31.7	27.9	31.7	29.3
36.2	30.9	28.5	34.1	30.8
34.5	27.6	27.6	32.4	28.9
30.7	29.5	29.0	32.0	31.4

1.) Calculate the mean for each car (to the nearest hundredth)

2.) Calculate the variance for each car (to the nearest tenth)

3.) Calculate the standard deviation for each car (to the nearest tenth)

	Mean	Variance	Stnd. Dev.
Hyundai Elantra			
Honda Civic			
Mazda 3			
Ford Focus			
Toyota Corolla			

Now that you have the variance and deviation of each car, answer the following questions

4.) Which car company was the most consistent with miles per gallon? How did you know?

5.) Would you use the mean or the standard deviation if you were trying to determine if the car companies were close to their predicted miles per gallon? Explain your reasoning