Name _____ Date _____

8A; Algebra 1 Period _____

Calculating Standard Deviation

Standard deviation is used to tell how far on average any data point is from the mean. The smaller the standard deviation, the closer the scores are on average to the mean. When the standard deviation is large, the scores are more widely spread out on average from the mean.

The **standard deviation** is the **square root** of the **variance**.

Variance measures how far a set of data is spread out. A variance of zero indicates that all of the data values are identical. All non-zero variances are positive. A small variance indicates that the data points tend to be very close to the mean, and to each other. A high variance indicates that the data points are very spread out from the mean, and from one another.

Variance is the **average of the squared distances** from each point to the mean.

Population vs. Sample Standard Deviation

When we are working with **every possible data point** of interest, we call this a **population** and use the population standard deviation, σ . When we have only a **sample** of all possible values we use the sample standard deviation, S. The formulas for these two differ very slightly, so their values tend to be slightly different. (see below)

What type of standard deviation should be used?

a) In a study of the heights of Merrick's teens, 100 students' heights were recorded.

b) In a study of the land areas of the states of the United States, the area of each state is used.

There will be a lot of new symbols used in this lesson. Here is a list of what some of them mean

$\sum$ = "the sum of ..."

n = number of pieces of data (population)

$n-1$ = number of pieces of data (sample)

$\overline{x}$ = mean (average) of data

x_i = each of the values in the data

$x_1, x_2, x_3, x_4, \ldots x_n$ (as i goes from 1 to n)

☆ Standard Deviation is better than MAD because it's more specific to population vs. sample.

☆ Standard Deviation is used for symmetrical data.

Practice Problem #1: Calculate the standard deviation of the following test data by hand. Use the chart below to record the steps.

Test Scores: 22, 99, 102, 33, 57, 75, 100, 81, 62, 29

Mean: 66 **n:** 10 (population (# of data values)) **n - 1:** 10 - 1 = 9 (sample)

Test Score (x)	Difference from the mean $(x - \bar{x})$	(Difference from the mean)^2 $(x - \bar{x})^2$
22	22 - 66 = -44	$(-44)^2 = 1936$
29	29 - 66 = -37	$(-37)^2 = 1369$
33	33 - 66 = -33	$(-33)^2 = 1089$
57	57 - 66 = -9	$(-9)^2 = 81$
62	62 - 66 = -4	$(-4)^2 = 16$
75	75 - 66 = 9	$(9)^2 = 81$
81	81 - 66 = 15	$(15)^2 = 225$
99	99 - 66 = 33	$(33)^2 = 1089$
100	100 - 66 = 34	$(34)^2 = 1156$
102	102 - 66 = 36	$(36)^2 = 1296$
	Sum of (Difference from the mean)^2 $\sum(x - \bar{x})^2$	+ 8338

Sum of (Difference from the Mean)^2 divided by n = $\frac{\sum(x-\bar{x})^2}{n} = \frac{8338}{10} = 833.8$ (This is Population Variance)

Sum of (Difference from the Mean)^2 divided by $(n-1)$ = $\frac{\sum(x-\bar{x})^2}{(n-1)} = \frac{8338}{9} = 926.4$ (This is Sample Variance)

Standard Deviation Squared

Final Step: Standard deviation = square root of what you just calculated (variance).

Population Standard deviation = $\sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{8338}{10}} = \sqrt{833.8} = 28.88$

Sample Standard deviation = $\sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}} = \sqrt{\frac{8338}{9}} = \sqrt{926.4} = 30.44$

Average distance from mean

Measures of Variation describe the dispersion or spread of a set of data. The VARIANCE and STANDARD DEVIATION describe how closely a set of data clusters about the mean. Referring back to the "Do Now", do you predict that Lisa or Melissa will have a higher variance value?

We use this formula: $\frac{\sum(x_i - \bar{x})^2}{n}$ or $\frac{\sum(x_i - \bar{x})^2}{n-1}$ to find the variance and we use this chart to help.

Calculator for standard deviation:
STAT | 1: Edit | put #'s in L_1
STAT → CALC | 1: 1-Var Stats | Calculate
* look for Sx (sample) or σx (population)

Lisa's data:

x_i	$\bar{x}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
50	84	50-84	$(-34)^2 = 1156$
98	84		
100	84		
94	84		
78	84		

Variance =

Standard Deviation =

Let's let the calculator do the work for us for Melissa's data! (85, 84, 83, 86, 82)

Calculator:

clear out L1 and L2
enter x_i data values into L1
STAT → CALC → 1-Var Stats L1

Standard deviation = _____

Variance = _____

The calculator will give us the standard deviation. So, how would we find the variance?

Calculator: Vars | 5: STATS | σx^2 (population) or Sx^2 (sample)

Examples:

1: Ten Merrick students are chosen at random to report their number of text messages in a half hour. Their numbers of text messages were: 15, 13, 12, 10, 9, 7, 5, 4, 3, 2

a) Determine if this data set is a population or a sample:

b) Find the mean:

c) Standard deviation:

d) Variance:

2: The following table shows English test scores for a whole class:

Grade	Frequency
95	4
85	13
75	11
70	6
65	2

a) Determine if this data set is a population or a sample:

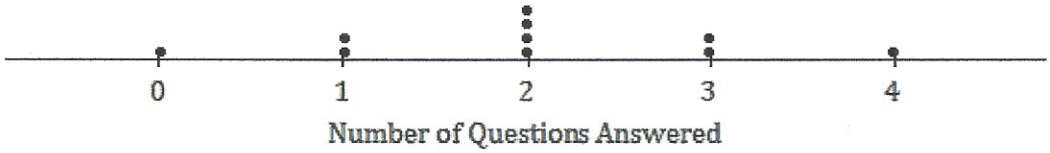
b) Find the mean:

c) Standard deviation to the nearest tenth:

d) Variance:

3. Ten people attended a talk at a conference. At the end of the talk, the attendees were given a questionnaire that consisted of four questions. The questions were optional, so it was possible that some attendees might answer none of the questions while others might answer 1, 2, 3, or all 4 of the questions (so the possible numbers of questions answered are 0, 1, 2, 3, and 4).

Suppose that the numbers of questions answered by each of the ten people were as shown in the dot plot below.



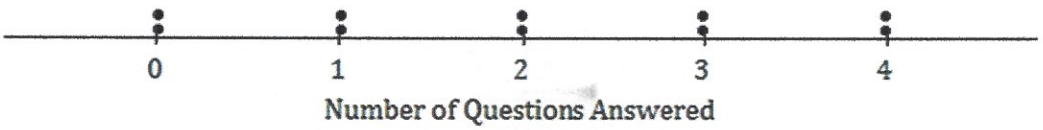
Number of Questions Answered

Use the statistical features of your calculator to find the mean and the standard deviation, to the nearest hundredth, of the data set. Would you use the population or sample standard distribution for this data?

Mean: _____

Standard Deviation: _____

4. Suppose the dot plot looked like this:

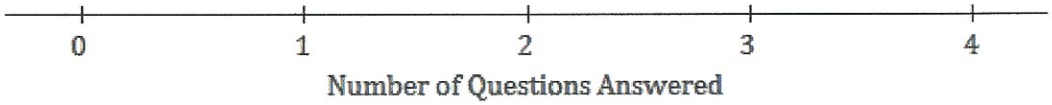


a. Use your calculator to find the mean and the standard deviation, to the nearest hundredth, of this distribution.

b. Remember that the size of the standard deviation is related to the size of the deviations from the mean. Explain why the standard deviation of this distribution is greater than the standard deviation in Example 3.

5. Suppose that every person answers all four questions on the questionnaire.

a. What would the dot plot look like?



b. What is the mean number of questions answered? (You should be able to answer without doing any calculations!)

c. What is the standard deviation? (Again, don't do any calculations!)