

Deviation and Standard Deviation

Activity

Part One. For a summer job, you were working in the quality control department for a computer company that manufactures computer parts. The specific part that you are to evaluate the quality of is supposed to be 8 micrometers in thickness. You obtained samples of four of these parts manufactured by the day shift and four parts manufactured by the night shift workers. Here are the findings:

Day Shift 7.9 8.0 8.2 8.3 Mean = _____ Median = _____

Night Shift 2 4 12 14 Mean = _____ Median = _____

1. Determine the mean and median for each shift and compare these with the desired level of 8 micrometers. Which shifts average was closer to the target? _____
2. If you needed to use one of these parts in your own computer on which shift would you prefer the part to have been manufactured? _____ Why?

Part Two.

3. In order to describe the variability of a set of data, a simple measure is to compute the **Range**. In statistical jargon, the Range is the gap between the largest and smallest value in the data set. Calculate the range for each shift.

Day shift: Range = _____ Night shift: Range = _____

4. A more sophisticated way of describing the variability of a data set is based on the notion of "**deviation from the mean**". For example: If the class average on a history test is 70, but you made a score of 87, your deviation would be +17. If your friend scored a 60 on the test, his or her deviation would be -10. Give the deviation from the mean for each part made on the night shift.

Data value: 2 4 12 14

Deviation: _____ _____ _____ _____

5. What properties do you notice about these deviations? And do you think these are true for all data sets, or just certain types? _____

6. The commonly used measure of variability is called the "**standard deviation**". Supposedly, the standard deviation tells us the size of a typical deviation from the mean for a particular set of data. Based on you answers to #4, guess at the value of the standard deviation. _____ Then take a guess as to how you think this value is might be calculated.

Part Three.

Here is how the standard deviation is actually calculated.

Step 1 - Find the mean

Step 2 - Find the deviation from the mean for each data value.

Step 3 - Square each of the deviations.

Step 4 - Find the sum of all of these squared deviations.

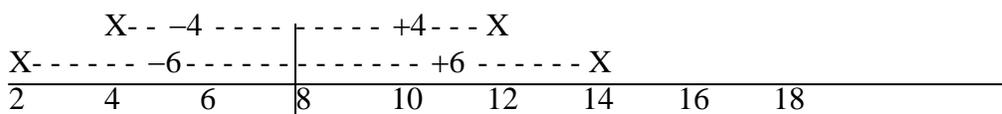
Step 5 - Average these by dividing by the number of observations.

Step 6 - Return to the units of the problem by taking the square root.

Example - Night shift data

	Data (x)	Deviation	Squared Dev.	
	2	- 6	36	Mean = $32/4 = 8.0$
	4	- 4	16	Sum of sq. dev. = 104
	12	+ 4	16	Ave. sq. dev. = $104/4 = 26$
	14	+6	36	
Totals	32	0	104	St. Dev. = $\sqrt{26} = 5.099$

Line plot:



7. Now try this with the following data: the number of minutes spent on the phone between 9:00 am and noon by six workers:

	Data (x)	Deviation	Squared Dev.	
	0			Mean =
	0			
	0			Sum of sq. dev. =
	2			
	5			Ave. sq. dev. =
	11			
Totals				St. Dev. =

8. Make a "line plot" of the data. Draw dotted horizontal lines showing the deviations from the mean.

9. In light of the current example, reconsider the response you gave to Question #5. Make up your own (small) set of data with a mean of 10. Then see if the deviations from the mean total to zero. Do you think this is true for all data sets?
