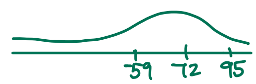
Learning Goal: Measures of Spread

Deviation is the difference between an individual value in a set of data and the mean for the data.

x = the individual value in a set
μ = the population mean
$\bar{x}$ = the sample mean

eg $=\{85, 95, 81, 60, 50, 40, 99\}$
$\bar{x} = 72\frac{6}{7}$



$s = \sqrt{\dfrac{(85-\bar{x})^2 + (95-\bar{x})^2 + (81-\bar{x})^2 + \cdots}{7-1}}$

$s = \sqrt{\dfrac{3154\frac{6}{7}}{6}}$

$s \doteq 22.931\%$

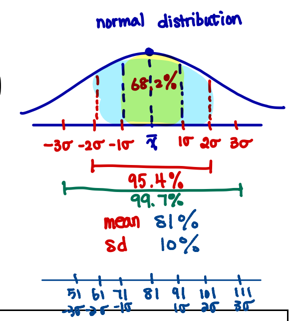
individual value

For the population, deviation is $x - \mu$ mean

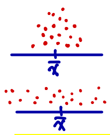
For a sample, deviation is $x - \bar{x}$

The larger the size of the deviations, the greater the spread in the data.

Class mean 81%
Std dev 10%



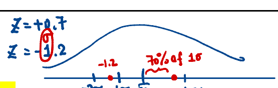
Standard Deviation – measure of variability (How much variation exists from the average)



- low standard deviation indicates the data points are very close to the mean
- High standard deviation indicates the data points are very far from the mean

Variance - how spread out numbers are

Measure	Sample	Population
Standard Deviation	$s = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}}$ = units n is the number of data in sample	$\sigma = \sqrt{\dfrac{\sum(x-\mu)^2}{N}}$ N is the number of data in population
Variance	$s^2 = \dfrac{\sum(x-\bar{x})^2}{n-1}$ = units²	$\sigma^2 = \dfrac{\sum(x-\mu)^2}{N}$



Z-Scores - the number of standard deviation that a datum is from the mean

Sample	Population
individual datum $z = \dfrac{x - \bar{x}}{s}$ mean, standard deviation s=standard deviation of the sample	$z = \dfrac{x-\mu}{\sigma}$ σ = standard deviation of the population

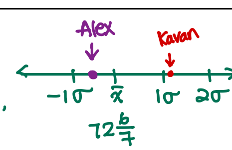
- Variables below mean have NEGATIVE Z-SCORES
- Variables above mean have POSITIVE Z-SCORES
- Variables equal to mean have ZERO Z-SCORE

eg. $\bar{x} = 72\frac{6}{7}$

Kavan's z-score $z = \dfrac{95 - 72\frac{6}{7}}{22.9305\ldots}$

$\doteq 1.040$

1σ above + 4% of a sd



Alex's $z = \dfrac{60 - 72\frac{6}{7}}{22.9305\ldots}$

$\doteq -0.486$

Ex. #1 -the number of hours five students spent studying statistics last week: 8 4 9 11 3

sample

a) Calculate Mean

$$\bar{x} = \frac{\sum x}{n} = \frac{8+4+9+11+3}{5}$$

$$= 7 \text{ hours}$$

b) Calculate Variance

$$s^2 = \frac{(x-\bar{x})^2}{n-1} = \frac{(8-7)^2+(4-7)^2+(9-7)^2+(11-7)^2+(3-7)^2}{5-1}$$

c) Calculate Standard Deviation $= 11.5 \text{ hours}^2$

$$\sqrt{s^2} = \sqrt{11.5}$$

$$\doteq 3.391 \text{ hours}$$

Ex. #2 – The number of summer jobs a sample of six students applied for: 17 15 23 7 9 13

Calculate the Standard Deviation

$$\text{mean} = \bar{x} = \frac{17+15+23+7+9+13}{6}$$

$$= 14 \text{ jobs}$$

$$s = \sqrt{\frac{\sum (x-14)^2}{6-1}}$$

$$\doteq 5.762 \text{ jobs}$$

Ex. #3 – Comparing Consistency of Two Types of Golf Clubs Consistency is the hallmark or a good golfer. Golf equipment manufacturers are constantly seeking ways to improve their products. Suppose that a recent golf innovation is designed to improve the consistency of its users. As a test a golfer was asked to hit 150 shots using a 7-iron, 75 of which were hit with his current club and 75 with the new innovation 7-iron. The distances were measured and recorded. Which 7-iron is more consistent?

Current	
Mean	150.55
Standard Deviation	5.79
Sample Variance	33.55

Innovative	
Mean	150.15
Standard Deviation	3.09
Sample Variance	9.56

Innovative is more consistent compared to the current model because the standard deviation and variance are smaller. It means the data is less spread out. Since the data is more clustered, the range of distance is smaller. It seems like the innovative golf club will consistently land balls within that range.

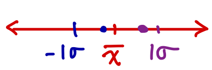
$z = \frac{x-\bar{x}}{s}$

Ex. #4 – If the mean is 68.1 and the standard deviation is 15.2, determine the z-scores for Audio Maximizer Ultra 3000 (score 67) and SchmederVox (score 75).

AMU3000:

$$z = \frac{67-68.1}{15.2}$$

$$\doteq -0.072$$



The AMU3000 has a z-score of -0.072, indicating that it is approximately 7% of one standard deviation below the mean.

SchmederVox:

$$z = \frac{75-68.1}{15.2}$$

$$\doteq 0.454$$

SchmederVox has a z-score of 0.454, indicating that it is approximately 45% of one standard deviation above the mean.