3.2 Measures of Variability (Spread)

**Statistics**
Values that describe a sample

- Standard deviation $\rightarrow S$
- Variance $\rightarrow S^2$
- Range $\rightarrow R$
- Interquartile range $\rightarrow IQR$

**Parameters**
Values that describe a population

- $\sigma$ "Sigma"
- $\sigma^2$
- $R$
- $IQR$
Measures of Spread: range, standard deviation, variance, IQR

* **range**: maximum - minimum

* **Standard deviation**: estimate the average distance each data value is from the mean

\[
S_x = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}} \quad \sigma_x = \sqrt{\frac{\sum (x-\mu)^2}{N}}
\]

* **Variance**: (standard deviation)^2

\[
S_x^2 = \frac{\sum (x-\bar{x})^2}{n-1} \quad \sigma_x^2 = \frac{\sum (x-\mu)^2}{N}
\]
Class A: 60, 70, 80, 90, 100
Class B: 60, 75, 80, 85, 100
Class C: 78, 79, 80, 81, 82, 85
Class D: 64, 64, 77, 77, 95, 100

$\bar{x} = 2.5$
Class A

<table>
<thead>
<tr>
<th>Data Value</th>
<th>Deviations from Mean</th>
<th>$L_1 = x - \bar{x}$</th>
<th>$L_2 = L_1^2$</th>
<th>$L_3 = L_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>-20</td>
<td>400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>-10</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>10</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>20</td>
<td>400</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\Sigma(x - x)^2 = \text{sum}(L_3) = 1000$

Variance $s^2 = \frac{\Sigma(x - \bar{x})^2}{n-1} = \frac{1000}{4} = 250$

Standard Deviation $s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}} = \sqrt{250} \approx 15.81$

Interpret the standard deviation:

On average, Class A test scores were about 15.81 pts from the mean.
### Class B

<table>
<thead>
<tr>
<th>$x$</th>
<th>$(x - \bar{x})$</th>
<th>$(x - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>-20</td>
<td>400</td>
</tr>
<tr>
<td>75</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>85</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>100</td>
<td>20</td>
<td>400</td>
</tr>
</tbody>
</table>

$\sum(x - \bar{x})^2 = 850$

Variance $= s^2 = \frac{\sum(x - \bar{x})^2}{n-1} = \frac{850}{4} = 212.5$

Standard Deviation $= s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}} = 14.58$

Interpret the standard deviation: On average, CLASS B test scores are about 14.58 pts from mean.
<table>
<thead>
<tr>
<th>Class</th>
<th>Range</th>
<th>Sample Mean ($\overline{x}$)</th>
<th>Sample Variance ($s^2$)</th>
<th>Sample Standard Deviation ($s$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>40</td>
<td>80</td>
<td>250</td>
<td>15.8</td>
</tr>
<tr>
<td>Class B</td>
<td>40</td>
<td>80</td>
<td>212.5</td>
<td>14.6</td>
</tr>
<tr>
<td>Class C</td>
<td>7</td>
<td>80.83</td>
<td>6.2</td>
<td>2.5</td>
</tr>
<tr>
<td>Class D</td>
<td>36</td>
<td>79.5</td>
<td>230.7</td>
<td>15.2</td>
</tr>
</tbody>
</table>
Normal Distribution (Gaussian or Bell-curve)

$z \rightarrow -3 \ -2 \ -1 \ 0 \ 1 \ 2 \ 3$

$\bar{x} \rightarrow \bar{x}-3s \ \bar{x}-2s \ \bar{x}-s \ \bar{x} \ \bar{x}+s \ \bar{x}+2s \ \bar{x}+3s$

- 68%
- 95%
- 99.7%
67% of the data is within 1 standard deviation of the mean (close to 68%)
95% of the data is within 2 standard deviations of the mean (exactly to 95%)
100% of the data is within 3 standard deviations of the mean (close to 99.7%)

Yes! The 2-dice sum data follows an approximately Normal distribution since it satisfies the 68-95-99.7 Rule!