Outline

1. Measuring Dispersion
2. Range
3. Standard Deviation
4. Chebychev’s Theorem
5. Conclusion
The Center isn’t Everything

Last time we looked at ways to measure the center of a set of data. While this is important, it is not the entire story.

Example

Give a lite chart and find the mean of each set of data.
The Center isn’t Everything

Last time we looked at ways to measure the center of a set of data. While this is important, it is not the entire story.

Example

Give a lite chart and find the mean of each set of data.

$$S_1 = \{55, 65, 70, 75, 85\}$$
Mean: $\bar{x}_1 = 70$
The Center isn’t Everything

Last time we looked at ways to measure the center of a set of data. While this is important, it is not the entire story.

Example

Give a lite chart and find the mean of each set of data.

\[ S_1 = \{55, 65, 70, 75, 85\} \]
Mean: \( \bar{x}_1 = 70 \)

\[ S_2 = \{67, 69, 71, 71, 72\} \]
Mean: \( \bar{x}_2 = 70 \)
Outline

1. Measuring Dispersion
2. Range
3. Standard Deviation
4. Chebychev’s Theorem
5. Conclusion
As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

**Range**

The range of a set of data is the difference between the highest and lowest values in the data set.

**Example**

Find the range of each of the data sets seen in the previous example.

1. \{55, 60, 70, 75, 85\}
2. \{67, 69, 71, 71, 72\}
As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

**Range**

The range of a set of data is the difference between the highest and lowest values in the data set.

**Example**

Find the range of each of the data sets seen in the previous example.

1. \{55, 65, 70, 75, 85\}
2. \{67, 69, 71, 71, 72\}
As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

**Range**

The range of a set of data is the difference between the highest and lowest values in the data set.

**Example**

Find the range of each of the data sets seen in the previous example.

1. \{55, 65, 70, 75, 85\}
2. \{67, 69, 71, 71, 72\}
Range

As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

The range of a set of data is the difference between the highest and lowest values in the data set.

Example

Find the range of each of the data sets seen in the previous example.

1. \(\{55, 65, 70, 75, 85\}\)
2. \(\{67, 69, 71, 71, 72\}\)
As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

**Range**

The range of a set of data is the difference between the highest and lowest values in the data set.

**Example**

Find the range of each of the data sets seen in the previous example.

1. \( \{55, 65, 70, 75, 85\} \)
   
   Range: \( 85 - 55 = 30 \)

2. \( \{67, 69, 71, 71, 72\} \)
As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

**Range**

The range of a set of data is the difference between the highest and lowest values in the data set.

**Example**

Find the range of each of the data sets seen in the previous example.

1. \{55, 65, 70, 75, 85\}
   
   \[
   \text{Range: } 85 - 55 = 30
   \]

2. \{67, 69, 71, 71, 72\}
As the previous example shows, we need to measure dispersion as well as center. Our first measure of dispersion is the range.

**Range**

The range of a set of data is the difference between the highest and lowest values in the data set.

**Example**

Find the range of each of the data sets seen in the previous example.

1. \{55, 65, 70, 75, 85\}
   
   Range: 85 - 55 = 30

2. \{67, 69, 71, 71, 72\}
   
   Range: 72 - 67 = 5
Unfortunately, the range is not enough to measure dispersion.

Example

Compute the mean and range of the data set

\[ S_3 = \{55, 57, 65, 65, 78, 85, 85\} \]

and draw a line chart.

\[ \bar{x}_3 = \frac{55 + 57 + 65 + 65 + 78 + 85 + 85}{7} = 70 \]

Range: 85 - 55 = 30
Unfortunately, the range is not enough to measure dispersion.

Example

Compute the mean and range of the data set

\[ S_3 = \{55, 57, 65, 65, 78, 85, 85\} \]

and draw a line chart.

\[
\bar{x}_3 = \frac{55 + 57 + 65 + 65 + 78 + 85 + 85}{7} = 70
\]

Range: \(85 - 55 = 30\)
Unfortunately, the range is not enough to measure dispersion.

**Example**

Compute the mean and range of the data set $S_3 = \{55, 57, 65, 65, 78, 85, 85\}$ and draw a line chart.

$$\bar{x}_3 = \frac{55 + 57 + 65 + 65 + 78 + 85 + 85}{7} = 70$$

Range: $85 - 55 = 30$
Range is Not Enough, Part I

Unfortunately, the range is not enough to measure dispersion.

Example

Compute the mean and range of the data set \( S_3 = \{55, 57, 65, 65, 78, 85, 85\} \) and draw a line chart.

\[
\bar{x}_3 = \frac{55 + 57 + 65 + 65 + 78 + 85 + 85}{7} = 70
\]

Range: \( 85 - 55 = 30 \)
The data in $S_3$ has the same mean and range as that in $S_2$, but $S_3$ is clearly more spread out, as seen below.
The data in $S_3$ has the same mean and range as that in $S_2$, but $S_3$ is clearly more spread out, as seen below.

Our next measure of dispersion is found by computing the distance between each data point and the mean.
<table>
<thead>
<tr>
<th>Outline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Measuring Dispersion</td>
</tr>
<tr>
<td>2. Range</td>
</tr>
<tr>
<td>3. <strong>Standard Deviation</strong></td>
</tr>
<tr>
<td>4. Chebychev’s Theorem</td>
</tr>
<tr>
<td>5. Conclusion</td>
</tr>
</tbody>
</table>
Variance

We start with the variance. There are two formulas for variance, depending on whether we are measuring an entire population or a sample of the population.

Computing the Variance for a Population
Let \( \{x_1, x_2, \ldots, x_N\} \) be data gathered from an entire population, and \( \mu \) the mean of the data. Then, the variance is:

\[
\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_N - \mu)^2}{N}
\]

Computing the Variance for a Sample
Let \( \{x_1, x_2, \ldots, x_n\} \) be a sample with mean \( \bar{x} \). The variance is:

\[
s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n - 1}
\]
Variance

We start with the variance. There are two formulas for variance, depending on whether we are measuring an entire population or a sample of the population.

**Computing the Variance for a Population**

Let \( \{x_1, x_2, \ldots, x_N\} \) be data gathered from an entire population, and \( \mu \) the mean of the data. Then, the variance is:

\[
\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_N - \mu)^2}{N}
\]

**Computing the Variance for a Sample**

Let \( \{x_1, x_2, \ldots, x_n\} \) be a sample with mean \( \bar{x} \). The variance is:

\[
s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n - 1}
\]
Variance

We start with the variance. There are two formulas for variance, depending on whether we are measuring an entire population or a sample of the population.

Computing the Variance for a Population
Let \( \{x_1, x_2, \ldots, x_N\} \) be data gathered from an entire population, and \( \mu \) the mean of the data. Then, the variance is:

\[
\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_N - \mu)^2}{N}
\]

Computing the Variance for a Sample
Let \( \{x_1, x_2, \ldots, x_n\} \) be a sample with mean \( \bar{x} \). The variance is:

\[
s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n - 1}
\]
Computing the Variance

Compute the variance of each sample.
Computing the Variance

Compute the variance of each sample.

\[ S_1 = \{55, 65, 70, 75, 85\} \]

\[ \bar{x}_1 = 70 \quad s_1^2 = 125 \]

<table>
<thead>
<tr>
<th>( x )</th>
<th>( x - \bar{x} )</th>
<th>( (x - \bar{x})^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>-15</td>
<td>225</td>
</tr>
<tr>
<td>65</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>85</td>
<td>15</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
</tr>
</tbody>
</table>
Computing the Variance

### Computing Variance

Compute the variance of each sample.

- **$S_1 = \{55, 65, 70, 75, 85\}$**
  - $\bar{x}_1 = 70$  
  - $s^2_1 = 125$

- **$S_3 = \{55, 57, 65, 65, 78, 85, 85\}$**
  - $\bar{x}_3 = 70$  
  - $s^2_3 = 159.7$

#### Sample $S_1$ Variance Calculation

<table>
<thead>
<tr>
<th>$x$</th>
<th>$x - \bar{x}$</th>
<th>$(x - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>-15</td>
<td>225</td>
</tr>
<tr>
<td>65</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>85</td>
<td>15</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>500</strong></td>
</tr>
</tbody>
</table>

#### Sample $S_3$ Variance Calculation

<table>
<thead>
<tr>
<th>$x$</th>
<th>$x - \bar{x}$</th>
<th>$(x - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>-15</td>
<td>225</td>
</tr>
<tr>
<td>57</td>
<td>-13</td>
<td>269</td>
</tr>
<tr>
<td>65</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>65</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>78</td>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>85</td>
<td>15</td>
<td>225</td>
</tr>
<tr>
<td>85</td>
<td>15</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>958</strong></td>
</tr>
</tbody>
</table>
Standard Deviation

When we compute variance, we square the differences so that our units come out squared as well. If we measure values in inches, the variance would be in square inches. To solve this problem, we the square root of the variance to get back to the correct units.

\[
\sigma = \sqrt{\sigma^2} \quad s = \sqrt{s^2}
\]

**Standard Deviation**

The standard deviation is found by taking the square root of the variance.

**Computing Standard Deviation**

Compute the standard deviation for \( S_1 \) and \( S_3 \) from the previous example.

\[
\sigma_1 = \sqrt{125} \approx 11.18 \\
\sigma_3 = \sqrt{159.7} \approx 12.64
\]
Standard Deviation

When we compute variance, we square the differences so that our units come out squared as well. If we measure values in inches, the variance would be in square inches. To solve this problem, we take the square root of the variance to get back to the correct units.

The standard deviation is found by taking the square root of the variance.

\[
\sigma = \sqrt{\sigma^2} \quad s = \sqrt{s^2}
\]

Computing Standard Deviation

Compute the standard deviation for \( S_1 \) and \( S_3 \) from the previous example.

\[
s_1 = \sqrt{125} \approx 11.18
\]
\[
s_3 = \sqrt{159.7} \approx 12.6
\]
When we compute variance, we square the differences so that our units come out squared as well. If we measure values in inches, the variance would be in square inches. To solve this problem, we take the square root of the variance to get back to the correct units.

**Standard Deviation**

The standard deviation is found by taking the square root of the variance.

\[ \sigma = \sqrt{\sigma^2} \quad s = \sqrt{s^2} \]

**Computing Standard Deviation**

Compute the standard deviation for \( S_1 \) and \( S_3 \) from the previous example.

1. \( s_1 = \sqrt{125} \approx 11.18 \)
2. \( s_3 = \sqrt{159.7} \approx 12.64 \)
Standard Deviation

When we compute variance, we square the differences so that our units come out squared as well. If we measure values in inches, the variance would be in square inches. To solve this problem, we take the square root of the variance to get back to the correct units.

\[ \sigma = \sqrt{\sigma^2} \quad s = \sqrt{s^2} \]

Computing Standard Deviation

Compute the standard deviation for \( S_1 \) and \( S_3 \) from the previous example.

1. \[ s_1 = \sqrt{125} \approx 11.18 \]
2. \[ s_3 = \sqrt{159.7} \approx 12.64 \]
When we compute variance, we square the differences so that our units come out squared as well. If we measure values in inches, the variance would be in square inches. To solve this problem, we take the square root of the variance to get back to the correct units.

The standard deviation is found by taking the square root of the variance.

\[ \sigma = \sqrt{\sigma^2} \quad s = \sqrt{s^2} \]

**Computing Standard Deviation**

Compute the standard deviation for \( S_1 \) and \( S_3 \) from the previous example.

1. \( s_1 = \sqrt{125} \approx 11.18 \)
2. \( s_3 = \sqrt{159.7} \approx 12.64 \)
Chebychev’s Theorem

If a distribution of numbers has a population mean $\mu$ and population standard deviation $\sigma$, the probability that a randomly chosen outcome has between $\mu - k$ and $\mu + k$ is at least $1 - \frac{\sigma^2}{k^2}$.

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.
Chebychev’s Theorem

Chebychev’s Theorem

If a distribution of numbers has a population mean $\mu$ and population standard deviation $\sigma$, the probability that a randomly chosen outcome has between $\mu - k$ and $\mu + k$ is at least $1 - \frac{\sigma^2}{k^2}$.

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.
Chebychev’s Theorem

If a distribution of numbers has a population mean $\mu$ and population standard deviation $\sigma$, the probability that a randomly chosen outcome has between $\mu - k$ and $\mu + k$ is at least $1 - \frac{\sigma^2}{k^2}$.

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.

$$k = \$1.00 \quad 1 - (8.50)^2 / 1 = -71.25$$
Measuring Dispersion

Range

Standard Deviation

Chebychev’s Theorem

Conclusion

Chebychev’s Theorem

If a distribution of numbers has a population mean $\mu$ and population standard deviation $\sigma$, the probability that a randomly chosen outcome has between $\mu - k$ and $\mu + k$ is at least $1 - \frac{\sigma^2}{k^2}$.

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.

\[ k = $1.00 \quad 1 - (8.50)^2/1 = -71.25 \]
\[ k = $8.50 \quad 1 - (8.50)^2/(8.50)^2 = 0.00 \]
Chebychev’s Theorem

Chebychev’s Theorem

If a distribution of numbers has a population mean \( \mu \) and population standard deviation \( \sigma \), the probability that a randomly chosen outcome has between \( \mu - k \) and \( \mu + k \) is at least \( 1 - \frac{\sigma^2}{k^2} \).

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.

\[
\begin{align*}
k &= $1.00 & 1 - \frac{(8.50)^2}{1} &= -71.25 \\
k &= $8.50 & 1 - \frac{(8.50)^2}{(8.50)^2} &= 0.00 \\
k &= $16.50 & 1 - \frac{(8.50)^2}{(16.50)^2} &= 0.73
\end{align*}
\]
Chebychev’s Theorem

If a distribution of numbers has a population mean $\mu$ and population standard deviation $\sigma$, the probability that a randomly chosen outcome has between $\mu - k$ and $\mu + k$ is at least $1 - \frac{\sigma^2}{k^2}$.

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.

$$k = \$1.00 \quad 1 - (8.50)^2/1 = -71.25$$
$$k = \$8.50 \quad 1 - (8.50)^2/(8.50)^2 = 0.00$$
$$k = \$16.50 \quad 1 - (8.50)^2/(16.50)^2 = 0.73$$
$$k = \$26.88 \quad 1 - (8.50)^2/(26.88)^2 = 0.90$$
Chebychev’s Theorem

If a distribution of numbers has a population mean $\mu$ and population standard deviation $\sigma$, the probability that a randomly chosen outcome has between $\mu - k$ and $\mu + k$ is at least $1 - \frac{\sigma^2}{k^2}$.

Example

The average order price at a department store is $51.25 with a standard deviation of $8.50. Find the smallest interval within which Chebychev’s theorem guarantees at least 90% of the sales fall.

\[
\begin{align*}
k &= $1.00 & 1 - \frac{(8.50)^2}{1} &= -71.25 \\
k &= $8.50 & 1 - \frac{(8.50)^2}{(8.50)^2} &= 0.00 \\
k &= $16.50 & 1 - \frac{(8.50)^2}{(16.50)^2} &= 0.73 \\
k &= $26.88 & 1 - \frac{(8.50)^2}{(26.88)^2} &= 0.90 \\
\end{align*}
\]

Between $24.37 and $78.13
Important Concepts

Things to Remember from Section 9-5

1. Computing the Range
2. Finding Variance
3. Finding Standard Deviation
4. Applying Chebychev’s Theorem
Important Concepts

Things to Remember from Section 9-5

1. Computing the Range
2. Finding Variance
3. Finding Standard Deviation
4. Applying Chebychev’s Theorem
Important Concepts

Things to Remember from Section 9-5

1. Computing the Range
2. Finding Variance
3. Finding Standard Deviation
4. Applying Chebychev’s Theorem
Important Concepts

Things to Remember from Section 9-5

1. Computing the Range
2. Finding Variance
3. Finding Standard Deviation
4. Applying Chebychev’s Theorem
Important Concepts

Things to Remember from Section 9-5

1. Computing the Range
2. Finding Variance
3. Finding Standard Deviation
4. Applying Chebychev’s Theorem
The last section in chapter 9 deals with the a shape of distribution which is very common in many different instances. This type of distribution is called a normal distribution.

For next time

- Read section 9-6
For next time

- Read section 9-6