Section 8.3: Variance and Standard Deviation

Example: Find the mean and median of these data sets.

A) 40, 40, 40, 100, 100, 100

\[ \text{Mean} = 70 \]
\[ \text{Median} = 70 \]

B) 60, 65, 65, 70, 75, 75, 80

\[ \text{Mean} = 70 \]
\[ \text{Median} = 70 \]

Definition: The standard deviation of a data set is the measure of how the data is spread about its mean.

The variance of a data set is the average of the square of the distance from the data value to the mean.

\[
\text{Notation:} \quad \begin{array}{ccc}
\text{Sample} & \text{Population} \\
S_x & \sigma \\
\bar{x} & \mu \\
\text{st. deviation} & \text{mean} \\
\end{array}
\]

\[
\text{Variance} = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_n - \mu)^2}{n}
\]

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

\[\text{St.dev} = \sqrt{\text{Var}}\]
\[(\text{St.dev})^2 = \text{Var}\]
Example: Compute the standard deviation and the variance.

<table>
<thead>
<tr>
<th>data</th>
<th>7</th>
<th>9</th>
<th>12</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>freq.</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

\[ \sigma^2 = 3.1423 \]

\[ \text{var} = (\sigma)^2 = (3.1423)^2 \]

Example: Compute the standard deviation and the variance. Let \( X \) be the number of Dr. Peppers drank during a semester.

<table>
<thead>
<tr>
<th>X</th>
<th>10</th>
<th>23</th>
<th>43</th>
<th>26</th>
<th>250</th>
</tr>
</thead>
<tbody>
<tr>
<td>freq.</td>
<td>50</td>
<td>30</td>
<td>49</td>
<td>73</td>
<td>1</td>
</tr>
</tbody>
</table>

\[ S_x = 19.55819 \]

\[ \text{var} = (S_x)^2 = (19.55819)^2 \]

Example: Compute the standard deviation.

<table>
<thead>
<tr>
<th>data</th>
<th>1</th>
<th>3</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>prob.</td>
<td>0.3</td>
<td>0.2</td>
<td>0.4</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Example: Let $X$ be a discrete random variable such that $0 \leq X \leq 90$. The random variable has an expected value of 15.3 and standard deviation of 3.5636.

A) What values of $X$ are within 1.25 standard deviations of the mean?

$\mu = 15.3$

$\mu \pm 1.25 \sigma$

$15.3 - 1.25(3.5636) = 10.4955$

$15.3 + 1.25(3.5636) = 19.7845$

Values of $X$ are 11, 12, ..., 19

B) What values of $X$ are up to 2 standard deviations above the mean?

$\mu = 15.3$

$\mu \pm 2\sigma$

$\mu + 2\sigma = 22.4272$

Answer: 0, 1, 2, 3, ..., 22

Chebyshev’s Inequality: Let $X$ be a random variable with mean $\mu$ and standard deviation $\sigma$. Then the probability that $X$ will be within $k$ standard deviations of the mean is

$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$

Example: The random variable $X$ has a mean of 50 and a standard deviation of 14. Estimate

$P(20 \leq X \leq 70)$.

$50 - k\sigma = 30 \Rightarrow 14k = 20 \Rightarrow k = \frac{10}{7}$

$P(30 \leq X \leq 70) \geq 1 - \frac{1}{\left(\frac{10}{7}\right)^2} = .51$
Example: The expected lifetime of a product is 2 years with a standard deviation of 3.5 months.
For a shipment of 5000 items, estimate the number of items that will last between 17.7 months and 30.3 months.

\[
P(17.7 \leq x \leq 30.3) = 1 - \frac{1}{(1.8)^2} = 0.6914
\]

\[
30.3 = 24 + k(3.5)
6.3 = 3.5k
k = \frac{6.3}{3.5} = 1.8
\]

\[
5000(0.6914) = 3457
\]

Answer: At least (2) 3457 items

\[
P(x < 17.7) \approx P(x > 30.3)
\]

Answer: \(1 - 0.6914\)