Chapter 14. Descriptive Statistics

Data set \{5, 3, 9, 7, 5, 7, 5, 8, 5\}

frequency table:

<table>
<thead>
<tr>
<th>Quiz score</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The size N of data set \{5, 3, 9, 7, 5, 7, 5, 8, 5\} is 9.

The average (or mean) is

\[
\frac{(5+3+9+7+5+7+5+8+5)}{9} = \frac{54}{9} = 6.
\]

Using the frequency table: the average is

\[
\frac{(3\times1 + 5\times4 + 7\times2 + 8\times1 + 9\times1)}{9} = \frac{3 + 20 +14 + 8 + 9}{9} = 6.
\]

The median of \{1, 3, 5, 9, 11\} is 5. First make sure that the numbers are arranged in increasing order. Then locate the point that separates the numbers into two equal parts.

\{1, 2, 4, 8, 9, 12\} has median = (4 + 8)/2 = 6.

Min, max and quartiles of \{1, 3, 4, 5, 5, 5, 6, 8, 9, 9, 11, 13, 13, 13, 13\}:

\[
\text{min} = 1, 3, 4 = Q_1, 5, 5, 5, 6 = M = Q_2, 8, 9, 9, 11 = Q_3, 13, 13, 13 = \text{max}
\]
Measures of spread:

\[ \text{range} = \text{max} - \text{min} = 13 - 1 = 12 \]

interquartile range \( = Q3 - Q1 = 11 - 4 = 7 \)

**standard deviation**: if the numbers in a data set are \( x_1, x_2, x_3, \ldots, x_N \) and the average is \( A = (x_1 + x_2 + x_3 + \ldots + x_N)/N \), then the \( \text{variance} = \text{var} = \)

\( \frac{(x_1 - A)^2 + (x_2 - A)^2 + (x_3 - A)^2 + \ldots + (x_N - A)^2)}{N} \)

and the **standard deviation** is the **square root** of the variance.

Example: data set 1 = \{2, 4, 5, 5, 5, 6, 6, 7\}

Average \( A = (2 + 4 + 5*4 + 6*2 + 7)/9 = 45/9 = 5 \)

<table>
<thead>
<tr>
<th>data</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>deviation from A</td>
<td>-3</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>squared deviations</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
Chapter 15. Chances, probabilities, and odds

**Sample space** of a random experiment **toss a coin three times** is:

\[ S = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\} . \]

**Multiplication rule:** count outcomes in two or more stages; suppose **in each stage the number of choices doesn't depend on choices made in other stages**; then the number of outcomes is the product of the number of choices in each stage.

\[
\text{Stage 1: President} \quad \times \quad \text{Stage 2: Vice President} \quad \times \quad \text{Stage 3: Secretary} \quad = \quad 60
\]

\[
\begin{array}{c}
\text{any of the 5 candidates} \\
5 \\
\end{array} 
\times 
\begin{array}{c}
\text{any of the 4 remaining candidates} \\
4 \\
\end{array} 
\times 
\begin{array}{c}
\text{any of the 3 remaining candidates} \\
3 \\
\end{array}
\]

### TABLE 15-1 Permutations and Combinations

<table>
<thead>
<tr>
<th>Notation</th>
<th>( nP_r )</th>
<th>( nC_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formula 1</strong></td>
<td>( nP_r = n(n-1)(n-2)\cdots(n-r+1) )</td>
<td>( nC_r = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!} )</td>
</tr>
<tr>
<td><strong>Formula 2</strong></td>
<td>( nP_r = \frac{n!}{(n-r)!} )</td>
<td>( nC_r = \frac{n!}{(n-r)!r!} )</td>
</tr>
<tr>
<td><strong>Applications</strong></td>
<td>Stud poker hands, rankings, committees with assignments</td>
<td>Draw poker hands, lottery tickets, coalitions, subsets</td>
</tr>
</tbody>
</table>
The number of 5 card hands (in which order matters) from 52 cards
\[ = {_{52}}P_{5} = 52 \times 51 \times 50 \times 49 \times 48 = 311,875,200 \]

and the number of unordered 5 card hands is
\[ = {_{52}}C_{5} = \frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2 \times 1} = \frac{311,875,200}{120} = 2,598,960 \]

If the sample space \( S \) is \( \{o_1, \ldots, o_N\} \), a probability assignment gives a probability \( Pr(o_i) \) to each outcome \( o_i \). Each probability is between 0 and 1, and \( Pr(o_1) + \ldots + Pr(o_N) = 1 \).

An event is any subset of the sample space. The probability of an event is the sum of the probabilities of the outcomes in that event.

Suppose we have a random experiment with sample space \( S = \{o_1, \ldots, o_N\} \). Suppose the outcomes are equally likely.

Then \( Pr(o_1) = \ldots = Pr(o_N) = 1/N. \)

Suppose an event \( E \) consists of \( K \) of these outcomes. Then \( Pr(E) = K/N. \)