Chapter 1: Overview

- Probability is a subarea of mathematics.
- Statistics is targeting for application.
- Drawing statistical conclusion should be based on probability theory.
- Probability uses assumptions.
- Statistics uses data.
- Statistical answer may be wrong.
- Probability answer is always correct if it has been proven.
Example 1.4.

- An article says “heart attack risk can be reduced by taking aspirin” (statement).

- Evidence: two group comparison.
  - Control group: 189 heart attacks in 11,034 people.
  - Aspirin group: 104 heart attacks in 11037 people.

- Q: does this support the statement.

- Answer: yes.

- Method: we will learn in Section 9.4.
Example


- Q: how to summarize the data?

- Answer:
  
  - Average:
    \[
    \bar{x} = \frac{1}{20}(612 + 623 + \cdots + 1201) = 964.95.
    \]

  - Sample variance:
    \[
    s^2 = \frac{1}{19}[(612 - 964.95)^2 + (623 - 964.95)^2
    \]
    \[
    + \cdots + (1201 - 964.95)^2]
    \]
    \[
    = 31790.26
    \]

  - Standard error:
    \[
    s = \sqrt{s^2} = \sqrt{31790.26} = 178.29.
    \]
Example

• A study about new medicine for leukemia patients.

• It randomly divided 42 patients into two groups:
  – Treatment: 21 patients taking medicine.
  – Control: 21 patients taking placebo.

• Drop off is allowed: death=1, drop off=0.

• Q: is the medicine effective?

• Answer: yes.

• The method will not be taught in this course.

• The data is behind.
<table>
<thead>
<tr>
<th>treat</th>
<th>drop</th>
<th>off</th>
<th>weeks</th>
<th>treat</th>
<th>drop</th>
<th>off</th>
<th>weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug</td>
<td>0</td>
<td>6</td>
<td>placebo</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>6</td>
<td>placebo</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>6</td>
<td>placebo</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>6</td>
<td>placebo</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>7</td>
<td>placebo</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>9</td>
<td>placebo</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>10</td>
<td>placebo</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>10</td>
<td>placebo</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>11</td>
<td>placebo</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>13</td>
<td>placebo</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>16</td>
<td>placebo</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>17</td>
<td>placebo</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>19</td>
<td>placebo</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>20</td>
<td>placebo</td>
<td>1</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>22</td>
<td>placebo</td>
<td>1</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>1</td>
<td>23</td>
<td>placebo</td>
<td>1</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>25</td>
<td>placebo</td>
<td>1</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>32</td>
<td>placebo</td>
<td>1</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>32</td>
<td>placebo</td>
<td>1</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>34</td>
<td>placebo</td>
<td>1</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drug</td>
<td>0</td>
<td>35</td>
<td>placebo</td>
<td>1</td>
<td>23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Statistics is very useful in

- social sciences,
- engineering,
- earth and atmospheric study,
- physics,
- political sciences,
- environmental sciences,
- education,
- public health,
- and etc.