This time intro: types of variables, samples & populations

next time: numerical & graphical descriptive methods

dichotomous

<table>
<thead>
<tr>
<th>Disease?</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>n = 93</td>
</tr>
<tr>
<td>No</td>
<td>n = 4</td>
</tr>
</tbody>
</table>

sample size

1 row for each subject

average

\[
\text{Mean} = \frac{\text{Sum of 15 & 09}}{n} = \frac{15 + 9}{9} = \frac{24}{9} = 2.67
\]
$$y_1 + y_2 + \ldots + y_n$$

mean = $\bar{y}$

$$\bar{y} = \frac{y_1 + y_2 + \ldots + y_n}{n} = \frac{93}{93} = 1$$

$$y_1 = 93$$

$$n = 93$$

Population mean

Sample mean

Sample size

0.7% of 93

Discrete data

The observed

Sample estimate of $\theta$

$\hat{\theta}$
mean = \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}

\begin{align*}
&= \frac{1}{n} \left( y_1 + y_2 + \ldots + y_n \right) \\
&= \frac{1}{n} \left( \sum_{i=1}^{n} y_i \right)
\end{align*}

\begin{align*}
\left( y_1 - \bar{y} \right) + \left( y_2 - \bar{y} \right) + \ldots + \left( y_n - \bar{y} \right) \\
= \sum_{i=1}^{n} (y_i - \bar{y})
\end{align*}
how do this? as possible as possible? wen wort is not
show is similar

so inter tested

we see R2

sample sample

factorial sample

mean = \frac{1}{5} M5

\frac{5}{1} (y_1 + y_2 + y_3 + y_4 + y_5)
at random with replacement = independent identical distributed (i.i.d.) sampling

\[
\begin{pmatrix}
\frac{1}{2} \\
\frac{9}{2}
\end{pmatrix}
\]

Sample

\[
\begin{pmatrix}
y_1 = 9 \\
y_2
\end{pmatrix}
\]

h = 2
at random
without replacement = simple random sampling (SRS)

more informative than IFD

<table>
<thead>
<tr>
<th>Variable</th>
<th>Possible Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>eye color</td>
<td>blue (1), brown (0)</td>
</tr>
<tr>
<td>hair color</td>
<td>black, brown, red, white</td>
</tr>
<tr>
<td>plant size</td>
<td>height (cm), # of leaves</td>
</tr>
<tr>
<td>blue</td>
<td>brown</td>
</tr>
</tbody>
</table>

quantitative