4.1 Samples {0.1} {0.2} {0.3} {0.4} {1.2} {1.3} {1.4} {2.3} {2.4} {3.4}

Means, $\bar{y}$ 0.5 1.0 1.5 2.0 2.5 3.0 3.5

where $\bar{y} = \frac{1}{n} \sum y_i$

The probability distribution of sample mean $\bar{y}$ is

<table>
<thead>
<tr>
<th>$\bar{y}$</th>
<th>0.5</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
<th>3.0</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(\bar{y})$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

$E(\bar{y}) = \sum \bar{y} p(\bar{y})$

$= (0.5)(.1) + (1.0)(.1) + 1.5(.2) + (2.0)(.2) + (2.5)(.2) + (3.0)(.1) + (3.5)(.1)$

$= 2$

$E(\bar{y}^2) = \sum \bar{y}^2 p(\bar{y})$

$= (0.5)^2 (.1) + (1.0)^2 (.1) + (1.5)^2 (.2) + (2.0)^2 (.2) + (2.5)^2 (.2) + (3.0)^2 (.1) + (3.5)^2 (.1) = 4.75$

$V(\bar{y}) = E(\bar{y}^2) - E(\bar{y})^2 = 4.75 - 4 = .75$

The probability distribution of $y$ is

<table>
<thead>
<tr>
<th>$y$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(y)$</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
</tr>
</tbody>
</table>

$E(y) = \sum y p(y) = 0(.2) + 1(.2) + 2(.2) + 3(.2) + 4(.2) = 2$

$E(y^2) = \sum y^2 p(y) = 0^2 (.2) + 1^2 (.2) + 2^2 (.2) + 3^2 (.2) + 4^2 (.2) = 6$

$\sigma^2 = V(y) = E(y^2) - E(y)^2 = 6 - 4 = 2$

So, $V(\bar{y}) = \frac{N-n}{N-1} \left( \frac{\sigma^2}{n} \right) = \frac{5-2}{5-1} \left( \frac{2}{2} \right) = \frac{3}{4} = .75$
4.2 Samples \{0.1\} \{0.2\} \{0.3\} \{0.4\} \{1.2\} \{1.3\} \{1.4\} \{2.3\} \{2.4\} \{3.4\} \\
\begin{array}{c|c|c|c|c|c|c|c|c|c|}
  s^2 & 0.5 & 2.0 & 4.5 & 8.0 & 0.5 & 2.0 & 4.5 & 0.5 & 2.0 \\
\end{array} \\
\text{where } s^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2 \\
The probability distribution of sample variance, \( s^2 \) is \\
\begin{array}{c|c|c|c|c|}
  s^2 & 0.5 & 2.0 & 4.5 & 8.0 \\
  p(s^2) & 0.4 & 0.3 & 0.2 & 0.1 \\
\end{array} \\
E(s^2) = \sum s^2 p(s^2) = 0.4(0.4) + 2.0(0.3) + 4.5(0.2) + 8.0(0.1) = 2.5 \\
\text{But, } E(s^2) = \frac{N}{N-1} \sigma^2 = \frac{5}{4} (2) = 2.5 \\
4.14 \hat{p} = \frac{\sum y_i}{n} = \frac{25}{30} = \frac{5}{6} = 0.83 \\
B = 2 \sqrt{\frac{\hat{p} \hat{q}}{N-1} \left( \frac{N-n}{N} \right)} = 2 \sqrt{\frac{(5/6)(1/6)}{29} \left( \frac{300-30}{300} \right)} = 0.131 \\
4.15 B = .05 \text{ } D = B^2 / 4 = (.05)^2 / 4 = .000625 \\
\text{From Equation (4.19), we have} \\
n = \frac{Npq}{(N-1)D + pq} = \frac{300 (5/6) (1/6)}{299 (.000625) + (5/6) (1/6)} = 127.90 \approx 128 \\
4.17 \hat{\tau} = N\bar{y} = 10000(12.5) = 125,000 \\
B = 2 \sqrt{N^2 \left( \frac{s^2}{n} \right) \left( \frac{N-n}{N} \right)} = 2 \sqrt{10000 \frac{1252}{100} \frac{10000-100}{10000}} = 70,412.50 \\
4.19 N = 1000, n = 10 \\
\bar{y} = \frac{\sum y_i}{n} = \frac{20}{10} = 2.0 \\
s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{\sum y_i^2 - n\bar{y}^2}{n-1} = \frac{60 - 10(4)}{9} = \frac{20}{9} = 2.22 \\
\hat{\mu} = \bar{y} = 2 \\
B = 2 \sqrt{\frac{s^2}{n} \left( \frac{N-n}{N} \right)} = 2 \sqrt{\frac{2.22}{10} \left( \frac{1000-10}{1000} \right)} = .938 \\
4.30 \text{ Choice D is the only correct answer, but should be modified to: It would be unlikely to get the observed sample proportion of 73% unless the actual percentage of all adults who want email service is between 69% and 77%.}
Ex 4.44 The proportion choosing “blamed the players” is dependent on the proportion choosing “blamed the owners” for the baseball strike. Thus, an approximate estimate of the true difference is

\[
(0.29 - 0.34) \pm 2 \sqrt{\frac{0.29 \cdot 0.71}{600 - 1} + \frac{0.34 \cdot 0.66}{600 - 1} + 2 \frac{0.29 \cdot 0.34}{600 - 1}}
\]

\[
\approx -0.05 \pm 0.06.
\]

So the confidence interval is (−0.11, 0.1), which contains 0. There is no significant evidence to suggest that the true proportions who blame the players and owners are really different. The observed difference could be simply due to chance.