Sample Size Requirement: Averages
The Soy Bean Oil Example

Use the information in the following setting to answer questions 1 through 3

An agronomist has designed a project to investigate the oil yield (in %) that can be expected from a new variety of soy bean. Specifically, the research team would like to estimate the average oil yield. In order for this estimate to be of any use the 95% confidence interval can be no wider than 1.0% (total).

A small pilot study has been done with this soy bean variety and the descriptive statistics from this are provided in Output 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SE Mean</th>
<th>StDev</th>
<th>Minimum</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>yield (%)</td>
<td>24</td>
<td>15.320</td>
<td>0.446</td>
<td>2.185</td>
<td>11.410</td>
<td>13.500</td>
<td>15.805</td>
<td>16.810</td>
<td>18.660</td>
</tr>
</tbody>
</table>

1) What is \( B \) the desired bound on the estimate for this setting:
\[
B = \frac{1.96}{2} = 0.98
\]

2) The z-value that will be used to compute the minimum necessary sample size is:
\[
\frac{1.96}{2} = 0.96
\]

3) The minimum necessary sample size required for the research team to meet their criteria is:
\[
n = \left(\frac{1.96 \cdot 2.185}{5}\right)^2 = 73.7
\]

So we need to sample at least 74

Required Sample Size– for estimating a population average
\[
n = \left(\frac{z \cdot S}{B}\right)^2
\]

Note: this formula is based on the assumption that the variable of interest is distributed normally (or approximately so).