(a) Interpret this interval in context of the data.

(b) Suppose another set of researchers reported a confidence interval with a larger margin of error based on the same sample of 1,155 Americans. How does their confidence level compare to the confidence level of the interval stated above?

(c) Suppose next year a new survey asking the same question is conducted, and this time the sample size is 2,500. Assuming that the population characteristics, with respect to how much time people spend relaxing after work, have not changed much within a year. How will the margin of error of the 95% confidence interval constructed based on data from the new survey compare to the margin of error of the interval stated above?

2 Thanksgiving spending, Part I The 2009 holiday retail season, which kicked off on November 27, 2009 (the day after Thanksgiving), had been marked by somewhat lower self-reported consumer spending than was seen during the comparable period in 2008. To get an estimate of consumer spending, 436 randomly sampled American adults were surveyed. Daily consumer spending for the six-day period after Thanksgiving, spanning the Black Friday weekend and Cyber Monday, averaged $84.71. A 95% confidence interval based on this sample is ($80.31, $89.11). Determine whether the following statements are true or false, and explain your reasoning.

(a) We are 95% confident that the average spending of these 436 American adults is between $80.31 and $89.11.

(b) This confidence interval is not valid since the distribution of spending in the sample is right skewed.

(c) 95% of random samples have a sample mean between $80.31 and $89.11.

(d) We are 95% confident that the average spending of all American adults is between $80.31 and $89.11.

(e) A 90% confidence interval would be narrower than the 95% confidence interval since we don’t need to be as sure about our estimate.

(f) In order to decrease the margin of error of a 95% confidence interval to a third of what it is now, we would need to use a sample 3 times larger.

(g) The margin of error is 4.4.
2 Sample Means

Does treatment using embryonic stem cells (ESCs) help improve heart function following a heart attack? Table 1 contains summary statistics for an experiment to test ESCs in sheep that had a heart attack. Each of these sheep was randomly assigned to the ESC or control group, and the change in their hearts’ pumping capacity was measured in the study. A positive value corresponds to increased pumping capacity, which generally suggests a stronger recovery. Our goal will be to identify a 95% confidence interval for the effect of ESCs on the change in heart pumping capacity relative to the control group.

A point estimate of the difference in the heart pumping variable can be found using the difference in the sample means:

$$\bar{x}_{esc} - \bar{x}_{control} = 3.50 - (-4.33) = 7.83$$

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>(\bar{x})</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESCs</td>
<td>9</td>
<td>3.50</td>
<td>5.17</td>
</tr>
<tr>
<td>control</td>
<td>9</td>
<td>-4.33</td>
<td>2.76</td>
</tr>
</tbody>
</table>

Table 1: Summary statistics of the embryonic stem cell study.

Can the \(t\)-distribution be used to make inference using the point estimate, \(\bar{x}_{esc} - \bar{x}_{control} = 7.83\)?

We check the two required conditions:

(i) Independence. Can we assume the conditions of the sheep in the study are independent of one another?

(ii) Normal. Are the distributions approximately normal?

Figure 1: Histograms for both the embryonic stem cell group and the control group. Higher values are associated with greater improvement. We don’t see any evidence of skew in these data; however, it is worth noting that skew would be difficult to detect with such a small sample.

We can quantify the variability in the point estimate, \(\bar{x}_{esc} - \bar{x}_{control}\), using the following formula for its standard error:

$$SE_{\bar{x}_{esc} - \bar{x}_{control}} = \sqrt{\frac{\sigma_{esc}^2}{n_{esc}} + \frac{\sigma_{control}^2}{n_{control}}}$$

We usually estimate this standard error using standard deviation estimates based on the samples:

$$SE_{\bar{x}_{esc} - \bar{x}_{control}} = \sqrt{\frac{s_{esc}^2}{n_{esc}} + \frac{s_{control}^2}{n_{control}}} 
\approx \sqrt{\frac{s_{esc}^2}{n_{esc}} + \frac{s_{control}^2}{n_{control}}} = \sqrt{\frac{5.17^2}{9} + \frac{2.76^2}{9}} = 1.95$$

Because we will use the \(t\)-distribution, we also must identify the appropriate degrees of freedom. This can be done using computer software. An alternative technique is to use the smaller of \(n_1 - 1\) and \(n_2 - 1\), which is the method we will typically apply in the examples and guided practice.
3 Calculate a 95% confidence interval for the effect of ESCs on the change in heart pumping capacity of sheep after they’ve suffered a heart attack.

A study of iron deficiency among infants compared samples of infants following different feeding regimens. One group contained breast-fed infants, while the children in another group were fed a standard baby formula without any iron supplements. A graphical display indicated that the blood hemoglobin levels in children (both breast-fed and formula-fed) are approximately normally distributed in each group. Here are the summary results on blood hemoglobin levels at 12 months of age:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample size</th>
<th>Sample mean</th>
<th>Sample SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast-fed</td>
<td>230</td>
<td>13.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Formula</td>
<td>230</td>
<td>12.4</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Construct and interpret (in context) a two sample confidence interval with 95% confidence.