Smokers  According to the National Center for Health Statistics (2004), 22.4% of adults are smokers. Suppose a random sample of 300 adults is obtained.

(a) Describe the sampling distribution of \( \hat{p} \), the sample proportion of adults who smoke.

The sample proportion \( \hat{p} \) is normally distributed with

\[
\mu_{\hat{p}} = 0.224 \quad \text{and} \quad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.224(1-0.224)}{300}} = 0.024.
\]

(b) In a random sample of 300 adults, what is the probability that at least 50 are smokers?

Let \( X \) be the number of adults out of 300 who are smokers.

\[
P(X \geq 50) = P\left(\frac{X}{300} \geq \frac{50}{300}\right)
= P(\hat{p} \geq 0.167)
= P(Z \geq \frac{0.167 - 0.224}{0.024})
= P(Z \geq -2.38)
= 1 - P(Z < -2.38)
= 1 - 0.0113
= 0.9887
\]

(c) Would it be unusual if a random sample of 300 adults results in 18% or less being smokers?

\[
P(\hat{p} \leq 0.18)
= P(Z \leq \frac{0.18 - 0.224}{0.024})
= P(Z \leq -1.83)
= 0.0336
\]

This would be an unusual outcome.
Page 402, Exercise 14

Teaching Supplies   According to the National Education Association, public school teachers spend an average of $443 of their own money each year to meet the needs of their students. Assume the standard deviation is $175. What is the probability that a random sample of 50 public school teachers spends an average of more than $400 each year to meet the needs of their students?

\[
P(\bar{x} > 400) = P\left(Z > \frac{400 - 443}{175/\sqrt{50}}\right) = P(Z > -1.74) = 1 - P(Z < -1.74) = 1 - 0.0409 = 0.9591
\]

Page 449, Exercise 4

A simple random sample of size \( n \) is drawn from a population that is known to be normally distributed. The sample mean \( \bar{x} \) is determined to be 104.3 and the sample standard deviation \( s \) is determined to be 15.9.

(a) Construct the 90\% confidence interval about the population mean if the sample size \( n = 15 \).

For the 90\% confidence interval \( \alpha = 0.10 \) which implied \( \alpha/2 = 0.05 \). Since the sample size is 15, the degrees of freedom is 14. According to Table V, \( t_{\alpha/2} = 1.761 \). The margin of error is

\[
E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 1.761 \frac{15.9}{\sqrt{15}} = 7.23
\]

Thus the 90\% confidence interval is

\[
(\bar{x} - E, \bar{x} + E) = (104.3 - 7.23, 104.3 + 7.23) = (97.07, 111.53).
\]

(b) Construct the 90\% confidence interval about the population mean if the sample size \( n = 25 \). How does increasing the sample size affect the width of the interval?

Since the sample size is 25, the degrees of freedom is 24. According to Table V, \( t_{\alpha/2} = 1.711 \). The margin of error is

\[
E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 1.711 \frac{15.9}{\sqrt{25}} = 5.44
\]

Thus the 90\% confidence interval is

\[
(\bar{x} - E, \bar{x} + E) = (104.3 - 5.44, 104.3 + 5.44) = (98.86, 109.74).
\]

The width of the interval has decreased.
(c) Construct the 95% confidence interval about the population mean if the sample size $n = 15$. Compare the results to those obtained in part (a). How does increasing the level of confidence affect the confidence interval?

For the 95% confidence interval $\alpha = 0.05$ which implied $\alpha/2 = 0.025$. Since the sample size is 15, the degrees of freedom is 14. According to Table V, $t_{\alpha/2} = 2.145$. The margin of error is

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 2.145 \frac{15.9}{\sqrt{15}} = 8.81.$$

Thus the 95% confidence interval is

$$(\bar{x} - E, \bar{x} + E) = (104.3 - 8.81, 104.3 + 8.81) = (95.49, 113.11).$$

The width of the interval has increased.

---

Page 450, Exercise 10

Math Achievement  

The following data represent the mathematics achievement test scores for a random sample of 15 male and 15 female students who had just completed high school in the United States, according to data obtained from the International Association for the Evaluation of Education Achievement study.

<table>
<thead>
<tr>
<th>Male</th>
<th>488</th>
<th>350</th>
<th>547</th>
<th>488</th>
<th>474</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>471</td>
<td>443</td>
<td>385</td>
<td>477</td>
<td>452</td>
</tr>
<tr>
<td></td>
<td>418</td>
<td>388</td>
<td>441</td>
<td>463</td>
<td>412</td>
</tr>
<tr>
<td>Female</td>
<td>433</td>
<td>389</td>
<td>520</td>
<td>479</td>
<td>454</td>
</tr>
<tr>
<td></td>
<td>563</td>
<td>411</td>
<td>458</td>
<td>398</td>
<td>337</td>
</tr>
<tr>
<td></td>
<td>418</td>
<td>492</td>
<td>442</td>
<td>494</td>
<td>514</td>
</tr>
</tbody>
</table>

(a) Verify that the scores for each gender are normally distributed with no outliers.

According to the following histograms and box-and-whiskers plots, the data are distributed approximately normally with no outliers.
(b) Obtain a point estimate for the population mean score of each gender.

\[ x_m = 446.5 \quad \text{and} \quad x_f = 453.5 \]

(c) Construct a 95% confidence interval for the population mean achievement score for males, assuming that \( \sigma = 64.8 \).

Since \( \sigma \) is known and the scores are normally distributed we will construct a Z-interval. For the 95% confidence interval \( \alpha = 0.05 \) which implied \( \alpha/2 = 0.025 \). According to Table IV, \( z_{\alpha/2} = 1.96 \). The margin of error is

\[ E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.96 \frac{64.8}{\sqrt{15}} = 32.8. \]

Thus the 95% confidence interval is

\[ (\bar{x} - E, \bar{x} + E) = (446.5 - 32.8, 446.5 + 32.8) = (413.7, 479.3). \]

(d) Construct a 95% confidence interval for the population mean achievement score for females, assuming that \( \sigma = 56.9 \).
Since $\sigma$ is known and the scores are normally distributed we will construct a $Z$-interval. For the 95% confidence interval $\alpha = 0.05$ which implied $\alpha/2 = 0.025$. According to Table IV, $z_{\alpha/2} = 1.96$. The margin of error is

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.96 \frac{56.9}{\sqrt{15}} = 28.8.$$  

Thus the 95% confidence interval is

$$ (\bar{x} - E, \bar{x} + E) = (453.5 - 28.8, 453.5 + 28.8) = (424.7, 482.3).$$

(e) Does there appear to be any difference between the scores of males and those of females? Since the 95% confidence intervals have a great deal of overlap, there does not appear to be a difference.

---

Page 451, Exercise 14

**Water Clarity**  
The campus at Joilet Junior College has a lake. A Secchi disk is used to measure the water clarity of the lake’s water by lowering the dish into the water and measuring the distance below water until the disk is no longer visible. The following measurements (in inches) were taken on the lake at various points in time over the course of a year.

82 64 62 66 68 43 38 26 68 56 54 66

(a) Use the data to compute a point estimate for the population mean and the population standard deviation.

$$\bar{x} = 57.8 \quad \text{and} \quad s = 15.4$$

(b) Because the sample size is small, we must verify that the data are normally distributed and do not contain any outliers. The figures below show the histogram and boxplot. Are the conditions for constructing a confidence interval about $\mu$ satisfied?
According to the histogram and the box-and-whiskers plot, the data are approximately normally distributed.

(c) Construct a 95% confidence interval for the mean Secchi disk measurement. Interpret this interval.

Since the data are approximately normally distributed but the population standard deviation is unknown, we will construct a $t$-interval. The sample size is $n = 12$ (degrees of freedom is 11), for the 95% confidence interval $\alpha = 0.05$ and $\alpha/2 = 0.025$ and so
from Table V, \( t_{\alpha/2} = 2.201 \). The margin of error is

\[ E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 2.201 \frac{15.4}{\sqrt{12}} = 9.8. \]

Thus the 95% confidence interval is

\[ (\bar{x} - E, \bar{x} + E) = (57.8 - 9.8, 57.8 + 9.8) = (48.0, 67.6). \]

If a large number of samples of size 12 are collected and the 95% confidence interval for each sample is computed, the true population mean will lie in 95% of the 95% confidence intervals.

(d) Construct a 99% confidence interval for the mean Secchi disk measurement. Interpret this interval.

Since the data are approximately normally distributed but the population standard deviation is unknown, we will construct a \( t \)-interval. The sample size is \( n = 12 \) (degrees of freedom is 11), for the 99% confidence interval \( \alpha = 0.01 \) and \( \alpha/2 = 0.005 \) and so from Table V, \( t_{\alpha/2} = 3.106 \). The margin of error is

\[ E = t_{\alpha/2} \frac{s}{\sqrt{n}} = 3.106 \frac{15.4}{\sqrt{12}} = 13.8. \]

Thus the 99% confidence interval is

\[ (\bar{x} - E, \bar{x} + E) = (57.8 - 13.8, 57.8 + 13.8) = (44.0, 71.6). \]

If a large number of samples of size 12 are collected and the 99% confidence interval for each sample is computed, the true population mean will lie in 99% of the 99% confidence intervals.

Page 451, Exercise 18

**Carbon Monoxide**  
From a random sample of 1201 Americans, it was discovered that 1139 of them lived in neighborhoods with acceptable levels of carbon monoxide (*Source: The Environmental Protection Agency*.)

(a) Obtain a point estimate for the proportion of Americans who live in neighborhoods with acceptable levels of carbon monoxide.

\[ \hat{p} = \frac{1139}{1201} = 0.948 \]
(b) Construct a 99% confidence interval for the proportion of Americans who live in neighborhoods with acceptable levels of carbon monoxide.

For the 99% confidence interval \( \alpha = 0.01 \) and \( \alpha/2 = 0.005 \) and \( z_{\alpha/2} = 2.576 \). The margin of error is

\[
E = z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 2.576 \sqrt{\frac{0.948(1 - 0.948)}{1201}} = 0.017.
\]

The 99% confidence interval is

\[
(\hat{p} - E, \hat{p} + E) = (0.948 - 0.017, 0.948 + 0.017) = (0.931, 0.965).
\]

(c) Suppose you wish to conduct your own study to determine the proportion of Americans who live in neighborhoods with acceptable levels of carbon monoxide. What sample size would be needed for the estimate to be within 1.5 percentage points with 90% confidence if you use the estimate obtained in part (a)?

For 90% confidence interval \( \alpha = 0.10 \) and \( \alpha/2 = 0.05 \) and thus \( z_{\alpha/2} = 1.645 \). Therefore the sample size, using the prior estimate of \( \hat{p} \) is

\[
n = \hat{p}(1 - \hat{p}) \left( \frac{z_{\alpha/2}}{E} \right)^2 = 0.948(1 - 0.948) \left( \frac{1.645}{0.015} \right)^2 = 592.9 \approx 593.
\]

(d) Suppose you wish to conduct your own study to determine the proportion of Americans who live in neighborhoods with acceptable levels of carbon monoxide. What sample size would be needed for the estimate to be within 1.5 percentage points with 90% confidence if you don’t have a prior estimate?

For 90% confidence interval \( \alpha = 0.10 \) and \( \alpha/2 = 0.05 \) and thus \( z_{\alpha/2} = 1.645 \). Therefore the sample size, with no prior estimate is

\[
n = 0.25 \left( \frac{z_{\alpha/2}}{E} \right)^2 = 0.25 \left( \frac{1.645}{0.015} \right)^2 = 3006.7 \approx 3007.
\]