8.4 Recall that if $Y_i$ is Exponential($\theta$) then $E(Y_i) = \theta$ and $V(Y_i) = \theta^2$. Hence we can use Theorem 5.12 to obtain

\[
E(\hat{\theta}_1) = E(\hat{\theta}_2) = E(\hat{\theta}_3) = E(\hat{\theta}_5) = \theta \\
V(\hat{\theta}_1) = \theta^2 \\
V(\hat{\theta}_2) = \frac{1}{4} (2\theta^2) = \frac{\theta^2}{2} \\
V(\hat{\theta}_3) = \frac{1}{9} (\theta^2 + 4\theta^2) = \frac{5\theta^2}{9} \\
V(\hat{\theta}_5) = \frac{1}{5} (3\theta^2) = \frac{3\theta^2}{5}
\]

The distribution of $\hat{\theta}_4$ can be obtained by using the methods of Section 6.6 in the text, with $F(y) = 1 - e^{-y/\theta}$. Then

\[
g_1(y) = \frac{1}{\theta} e^{-y/\theta} (e^{-y/\theta})^2 = \frac{1}{\theta} e^{-3y/\theta}
\]

which is an exponential distribution with mean $\frac{1}{3}$.

\[
E(\hat{\theta}_4) = \frac{1}{3} \\
V(\hat{\theta}_4) = \frac{\theta^2}{9}
\]

a. The unbiased estimators are $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$, and $\hat{\theta}_5$.

b. Among these four estimators, $\hat{\theta}_5 = \bar{Y}$ has the smallest variance.

8.6

a. For the Poisson distribution, $E(Y_i) = \lambda$ and $E(\bar{Y}) = \lambda$. Hence $\hat{\lambda} = \bar{Y}$ is an unbiased estimator for $\lambda$.

b. In order to find $E(\bar{Y}^2)$, use the fact that $V(Y) = \lambda$ and $E(Y^2) = V(Y) + [E(Y)]^2 = \lambda + \lambda^2$. Then $E(C) = 3E(Y) + E(\bar{Y}^2) = 4\lambda + \lambda^2$.

c. Since $E(\bar{Y}) = \lambda$, $E(\bar{Y}^2) = V(\bar{Y}) + [E(\bar{Y})]^2 = \frac{\lambda}{n} + \lambda^2$, we construct an estimator $\bar{\theta} = \bar{Y}^2 + \bar{Y} (4 - \frac{1}{n})$. Considering

\[
E(\bar{\theta}) = \frac{\lambda}{n} + \lambda^2 + 4\lambda - \left(\frac{1}{n}\right) \lambda = 4\lambda + \lambda^2.
\]

Thus, $\bar{\theta}$ is an unbiased estimator of $E(C)$.

8.8

a. For the uniform distribution given here, $E(Y_i) = \theta + \frac{1}{2}$. Hence $E(\bar{Y}) = \theta + \frac{1}{2}$ and the bias is $B = E(\bar{Y}) - \theta = \frac{1}{2}$.

b. An unbiased estimator of $\theta$ can be constructed by using $\bar{\theta} = \bar{Y} - \frac{1}{2}$, which has

\[
E(\bar{\theta}) = \theta.
\]

c. If $\bar{Y}$ is used as an estimator, then

\[
V(\bar{Y}) = \frac{V(Y)}{n} = \frac{1}{12n} \quad \text{and} \quad \text{MSE} = V(\bar{Y}) + B^2 = \frac{1}{12n} + \frac{1}{4}.
\]

8.18 The point estimate of $\mu$ is $\bar{Y} = 7.2\%$, and the bound on the error of estimation is $2\sigma_y$.
With $n = 200$ and $s = 5.6\%$, we have

\[
2\sigma_y = 2 \frac{s}{\sqrt{n}} \approx 2 \frac{5}{\sqrt{200}} = \frac{2(5.6)}{\sqrt{200}} = .79
\]

8.20 The value .54 is a point estimate of $p$. A two-standard-deviation bound on the error of estimation is

\[
2\sqrt{\frac{.54(1-.54)}{n}} = 2\sqrt{\frac{.294}{n}} = 2\sqrt{\frac{(1.54)(.46)}{1000}} = .03
\]

Note that $\bar{Y} = .51$. Thus we can conclude that a majority of individuals in this age group feel that religion is a very important part of their lives.
8.22 The point estimate for $p$ is $\hat{p} = \frac{3}{5}$. The bound on the error of estimation is
\[
2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2\sqrt{\frac{(\frac{3}{5})(\frac{2}{5})}{192}} = .023
\]

8.36 Use the fact that $Z = \frac{Y - \mu}{\sigma} = Y - \mu$ has a standard normal distribution.
   a. The 95% confidence interval for $\mu$ is $(Y - 1.96, Y + 1.96)$ since
   \[
P(-1.96 \leq Z \leq 1.96) = .95
   \]
   \[
P(-1.96 \leq Y - \mu \leq 1.96) = .95
   \]
   \[
P(Y - 1.96 \leq \mu \leq Y + 1.96) = .95
   \]
   b. Since
   \[
P(Z \leq -1.645) = .05
   \]
   \[
P(Y - \mu \leq -1.645) = .05
   \]
   \[
P(\mu \geq Y + 1.645) = .05
   \]
   Hence $Y + 1.645$ is the 95% upper limit for $\mu$.
   c. Similarly, $Y - 1.645$ is the 95% lower limit for $\mu$.

8.42 a. $\hat{p} = \frac{268}{500} = .536$. Therefore, an approximate 98% confidence interval for $p$ is
\[
\hat{p} \pm z_{.01}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .536 \pm 2.33\sqrt{\frac{(.536)(.464)}{500}} = .536 \pm .052 \text{ or } (.484, .588).
\]
b. Since the interval does include $p = .51$, we cannot conclude that there is a difference in the graduation rates before and after Proposition 48.

8.44 The parameter to be estimated in this exercise is $\mu$, the average number of days required for treatment of patients. The 95% confidence interval is approximately
\[
\bar{y} \pm z_{.025\sqrt{n}} \text{ or } 5.4 \pm 1.96\left(\frac{3.1}{\sqrt{500}}\right) \text{ or } 5.4 \pm .27 \text{ or } (5.13, 5.67)
\]

8.70 a. $n = 20$, $\bar{x} = 419$, $s = 57$. Then the 90% confidence interval for the mean SAT scores for urban high school seniors is
\[
\bar{y} \pm t_{.05\sqrt{n}}\left(\frac{s}{\sqrt{n}}\right)
\]
where $t_{.05}$ is based on $n - 1 = 19$ degrees of freedom. From the Appendix, this is $t_{.05} = 1.729$. Then the confidence interval is
\[
419 \pm 1.729\left(\frac{57}{\sqrt{50}}\right) = 419 \pm 22.04 = (396.96, 441.04).
\]
b. The interval does include 422. Thus 422 is a believable value for $\mu$ at the 90% confidence level. However, numbers such as 397, 410, and 441, for example, are also believable values for $\mu$.

c. Given $n = 20$, $\bar{x} = 455$, $s = 69$, the 90% confidence interval for the mean mathematics SAT score is
\[
\bar{y} \pm t_{.05\sqrt{n}}\left(\frac{s}{\sqrt{n}}\right) = 455 \pm 1.729\left(\frac{69}{\sqrt{20}}\right) = 455 \pm 26.67 = (428.33, 481.67).
\]The interval does include 474. We would conclude, based on our 90% confidence interval, that the true mean mathematics SAT score is not different from 474.

8.74 For the $n = 12$ measurements given here, calculate $\Sigma y_i = 108$ and $\Sigma y_i^2 = 1426$. Then
\[
\bar{y} = \frac{\Sigma y_i}{12} = 9 \quad \text{and} \quad s^2 = \frac{\Sigma y_i^2 - (\Sigma y_i)^2}{11} = 41.2727
\]
The 90% confidence interval is then
\[
\bar{y} \pm t_{.05\sqrt{n}}\left(\frac{s}{\sqrt{n}}\right) = \pm 1.796\sqrt{\frac{41.2727}{12}} = 9 \pm 3.33 \text{ or } (5.67, 12.33).
\]
8.76 a. Let $\mu_1 = \text{mean verbal score for engineering students}$ and $\mu_2 = \text{mean verbal score for language/literature students}$. Then the 95% confidence interval is

$$(\bar{y}_1 - \bar{y}_2) \pm t_{.025}\sqrt{\frac{S_p^2}{n_1} + \frac{1}{n_2}}$$

where $t_{.025} = 2.048$ with 28 degrees of freedom. Then,

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} = \frac{14(42)^2 + 14(45)^2}{28} = 1894.5$$

and the confidence interval is

$$446 - 534 \pm 2.048\sqrt{1894.5\left(\frac{1}{15} + \frac{1}{15}\right)} = -88 \pm 32.55 = (-120.55, -55.45)$$

b. Similar to part a. Let $\mu_1 = \text{mean math score for engineering students}$ and $\mu_2 = \text{mean math score for language/literature students}$. First, calculate

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} = \frac{14(57)^2 + 14(52)^2}{28} = 2976.5$$

Then the interval is

$$548 - 517 \pm 2.048\sqrt{2976.5\left(\frac{1}{15} + \frac{1}{15}\right)} = 31 \pm 40.80 = (-9.80, 71.80)$$

c. The 95% confidence intervals indicate that a significant difference exists in the mean verbal scores for students in engineering and language/literature (since both endpoints of the interval are negative). However, the other interval does not indicate that a significant difference exists in the mean math scores for students in engineering and language/literature, since 0 is in the interval.

d. We assume that the verbal (math) scores for the two groups are randomly and independently selected from two normal distributions with common variance.