1. For each of the following, determine the values of $\mu_\sigma$ and $\sigma_\sigma$ for the given population characteristics and sample sizes. Also state whether or not we can assume that the distribution of the sample mean is approximately normal. (6 points each)
   (a) $\mu = 80, \sigma = 14, n = 49$
   (b) $\mu = 64, \sigma = 18, n = 36$
   (c) $\mu = 52, \sigma = 10, n = 21$
   (d) $\mu = 27, \sigma = 6, n = 15$ and the population is approximately normal

2. The most famous geyser in the world, Old Faithful in Yellowstone National Park, has a mean time between eruptions of 85 minutes. If the interval of time between eruptions is normally distributed with a standard deviation of 21.25 minutes, answer the following. (5 points each)
   (a) What is the probability that a random sample of 10 time intervals between eruptions has a mean longer than 95 minutes?
   (b) What is the probability that a random sample of 20 time intervals between eruptions has a mean longer than 95 minutes?
   (c) What is the probability that a random sample of 30 time intervals between eruptions has a mean longer than 95 minutes?
   (d) What is happening to this probability as the sample size increases? How do you account for this happening?

3. For each of the following, determine the values of $\mu_\pi$ and $\sigma_\pi$ for the given population characteristics and sample sizes. Also state whether or not we can assume that the distribution of the sample proportion is approximately normal. (6 points each)
   (a) $\pi = 0.65, n = 10$
   (b) $\pi = 0.72, n = 32$
   (c) $\pi = 0.58, n = 25$

4. A report released in May 2005 by First Data Corp. indicated that 43% of adults had received a “phishing” contact (a bogus email that replicates an authentic site for the purpose of stealing personal information). Suppose a random sample of 800 adults is obtained. (5 points each)
   (a) What is the probability that no more than 40% of the sample have received a “phishing” contact?
   (b) Would it be unusual if in this sample 48% or more had received “phishing” contacts?

5. A sociologist wishes to estimate the percentage of Americans who favor affirmative action programs for women and minorities for admission to colleges and universities. What sample size should be obtained if she wishes the estimate to be within 3 percentage points with 98% confidence? (8 points)
6. In a Harris Poll conducted October 20 – 25, 2004, 381 of 2114 randomly selected adults who follow professional football said that the Green Bay Packers were their favorite team.  

(a) Construct a 90% confidence interval for the proportion of all adults who follow professional football who say the Green Bay Packers is their favorite team.  

(b) Construct a 98% confidence interval for the proportion of all adults who follow professional football who say the Green Bay Packers is their favorite team.  

(c) What effect did increasing the confidence level have on the width of the interval? Should this be expected to occur whenever we increase the confidence level? Why?

7. A Gallup poll conducted January 23, 2003 – February 10, 2003, asked American teens (age 13 – 17) how much time the spent each week using the Internet. Initial survey results support that $\sigma = 6.6$ hours. How many subjects are needed to estimate the average time American teens spend on the Internet each week within half an hour with 97.5% confidence?  

8. Based on a random sample of 1120 Americans 15 years of age or older, the mean amount of sleep per night is 8.17 hours. Assume the population standard deviation for amount of sleep per night is 1.2 hours.  

(a) Give a 94% confidence interval for the mean amount of sleep per night for all Americans age 15 years or older.  

(b) Give a 99% confidence interval for the mean amount of sleep per night for all Americans age 15 years or older.

9. The following data represent the concentration of organic carbon (mg/L) collected from organic soil. Construct a 99% confidence interval for the mean concentration of dissolved organic carbon collected from organic soil.  

<table>
<thead>
<tr>
<th>22.74</th>
<th>29.8</th>
<th>27.1</th>
<th>16.51</th>
<th>6.51</th>
<th>8.81</th>
<th>5.29</th>
<th>20.46</th>
<th>14.9</th>
<th>33.67</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.91</td>
<td>14.86</td>
<td>15.91</td>
<td>15.35</td>
<td>9.72</td>
<td>19.8</td>
<td>14.86</td>
<td>8.09</td>
<td>17.9</td>
<td>18.3</td>
</tr>
<tr>
<td>5.2</td>
<td>11.9</td>
<td>14</td>
<td>7.4</td>
<td>17.5</td>
<td>10.3</td>
<td>11.4</td>
<td>5.3</td>
<td>15.72</td>
<td>20.46</td>
</tr>
<tr>
<td>16.87</td>
<td>15.42</td>
<td>22.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. Severe Acute Respiratory Syndrome (SARS) is a viral respiratory illness. It has the distinction of being the first new communicable disease of the 21st century. Researchers wanted to know the incubation period of patients with SARS, so they randomly selected 81 SARS patients and found, for these patients, the mean incubation period was 4.6 days with a standard deviation of 15.9 days. Give a 98.5% confidence interval for the mean incubation period for the SARS virus.  

(6 points)
1. (a) $\mu_x = \mu = 80, \sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{14}{7} = 2$, and since the sample size is more than 30, the sample mean distribution is approximately normal

(b) $\mu_x = \mu = 64, \sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{18}{6} = 3$, and since the sample size is more than 30, the sample mean distribution is approximately normal

(c) $\mu_x = \mu = 52, \sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{21}} \approx 2.182$, and since the sample size is less than 30 and we know nothing of the shape of the population, the sample mean distribution cannot be assumed to be approximately normal

(d) $\mu_x = \mu = 27, \sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{15}} \approx 1.549$, and since the population is approximately normal, the sample mean distribution is approximately normal even though the sample size is less than 30

2. Note that since the time between eruptions is normally distributed, the sample mean distribution can be considered approximately normal regardless of sample size.

(a) $P(\bar{x} \geq 95) = \text{normalcdf}(95, \infty, 85, \frac{21.25}{\sqrt{10}}) \approx 0.0684$

(b) $P(\bar{x} \geq 95) = \text{normalcdf}(95, \infty, 85, \frac{21.25}{\sqrt{20}}) \approx 0.0177$

(c) $P(\bar{x} \geq 95) = \text{normalcdf}(95, \infty, 85, \frac{21.25}{\sqrt{30}}) \approx 0.00498$

(d) As the sample size increases, the probability being considered is getting lower. This makes sense because as the sample size increases, the standard deviation of the sample mean distribution gets smaller – resulting in less chance of moving far away from the mean.

3. (a) $\mu_p = \pi = 0.65, \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.65(0.35)}{10}} \approx 0.1508$

   $n\pi = 10(0.65) = 6.5$ and $n(1-\pi) = 10(0.35) = 3.5$ Since these are not both 10 or more, we cannot assume the distribution of the sample proportion is normal.

(b) $\mu_p = \pi = 0.72, \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.72(0.28)}{32}} \approx 0.0794$

   $n\pi = 32(0.72) = 23.04$ and $n(1-\pi) = 32(0.28) = 8.96$ Since these are not both 10 or more, we cannot assume the distribution of the sample proportion is normal.

(c) $\mu_p = \pi = 0.58, \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.58(0.42)}{25}} \approx 0.0987$

   $n\pi = 25(0.58) = 14.5$ and $n(1-\pi) = 25(0.42) = 10.5$ Since these are both 10 or more, we can assume the distribution of the sample proportion is normal.
4. Note that since \( n\pi = 800(0.43) = 344 \) and \( n(1 - \pi) = 800(0.57) = 456 \) are both greater than 10, we may assume that the distribution of the sample proportion is approximately normal.

   (a) \( P(p \leq 0.40) = \text{normalcdf}\left(-\infty, 0.40, 0.43, \sqrt{\frac{0.43(0.57)}{800}}\right) = 0.04323 \)

   (b) \( P(p \geq 0.48) = \text{normalcdf}\left(0.48, \infty, 0.43, \sqrt{\frac{0.43(0.57)}{800}}\right) = 0.00214 \), meaning that only about 0.2% of the time would we see such a sample. So it would be considered unusual.

5. \( n = 0.25\left(\frac{z^*}{B}\right)^2 \)

   \( z^* = \text{invNorm}(0.99) = 2.3263 \) and \( B \) is desired to be 0.03

   So \( n = 0.25\left(\frac{2.3263}{0.03}\right)^2 \approx 1503.24 \), meaning we would need **1504 subjects**

6. Note that the sample size is so large that \( n\pi = 381 \) and \( n(1 - \pi) = 1733 \). Thus the method of constructing confidence intervals we have discussed is valid.

   (a) **1-PropZInt** \((0.16648, 0.19398)\)

   (b) **1-PropZInt** \((0.16078, 0.19968)\)

   (c) The interval got wider when the confidence level was raised. This should be expected to happen in general because we are wanting to have more confidence in our answer so we are forced to allow for a larger range of values for the sample proportion.

7. \( n = \left(\frac{z^*\sigma}{B}\right)^2 \)

   \( z^* = \text{invNorm}(0.9875) = 2.2414 \)

   We are wanting an error of no more than \( B = \frac{1}{2} \), and we are told \( \sigma = 6.6 \).

   So, \( n = \left(\frac{2.2414(6.6)}{0.5}\right)^2 \approx 875.36 \), so we would need 876 subjects

8. Note that since the sample size is more than 30, the distribution of the sample mean is approximately normal, so the technique for building confidence intervals we have discussed is valid.

   (a) **ZInterval** \((8.1026, 8.2374)\) hours per night

   (b) **ZInterval** \((8.0776, 8.2624)\) hours per night
9. Since we do not know the population’s standard deviation, we must construct a T-Interval, and since the sample size is more than 30, the process is valid.

\[ T \text{Interval } (12.403, 19.443) \text{ mg/L} \]

So the mean value of organic carbon in organic soil is between 12.403mg/L and 19.443mg/L.

10. Since we do not know the population’s standard deviation, we must construct a T-Interval, and since the sample size is more than 30, the process is valid.

\[ T \text{Interval } (0.20798, 8.992) \text{ days} \]