1. Suppose that the diameter of a population of Douglas fir trees in the Pacific Northwest is normally distributed with mean 50 cm and standard deviation 12 cm.
   a. What percentage of trees in the population have a diameter more than 75 cm?
   b. A certain tree’s diameter is in the 33rd percentile. What is the diameter?
   c. If 25 trees were randomly sampled from this population, what is the probability that their average diameter is less than 45 cm?

2. Two dietary instruments that are often used in studies to assess consumption of specific foods are the food-frequency questionnaire (FFQ) and the dietary record (DR). The major difference between the two methods is that the FFQ collects records of certain foods eaten over a certain period of time and then nutrient intake is estimated, while the DR calculates nutrient intake exactly from a record of every food ingested. The FFQ less expensive to administer but is considered less accurate than the DR (the gold standard of dietary analyses). To validate the FFQ, 173 nurses participated in a study to record their food intake using both methods and total caloric intake (per day) was obtained for the two methods. Descriptive statistics for each method can be found below (calor_dr for the DR and calor_ffq for the FFQ). Suppose that the data do not come from a normal distribution.

   Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>TrMean</th>
<th>StDev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>calor_dr</td>
<td>173</td>
<td>1619.9</td>
<td>1606.0</td>
<td>1613.3</td>
<td>323.4</td>
<td>24.6</td>
</tr>
<tr>
<td>calor_ffq</td>
<td>173</td>
<td>1371.7</td>
<td>1297.6</td>
<td>1344.3</td>
<td>482.1</td>
<td>36.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>calor_dr</td>
<td>910.0</td>
<td>2518.0</td>
<td>1410.0</td>
<td>1784.5</td>
</tr>
<tr>
<td>calor_ffq</td>
<td>463.2</td>
<td>3077.3</td>
<td>1031.2</td>
<td>1591.0</td>
</tr>
</tbody>
</table>

   a. Calculate a 95% confidence interval for the true mean caloric intake for the FFQ method.
   b. It was mentioned above that the data do not come from a normal distribution, would you still consider the confidence interval that you calculated above valid? Why or why not?
   c. (TRUE or FALSE and say why) We are 95% confident that the average caloric intake of the 173 nurses using the FFQ method is between (interval calculated in part b).
   d. Suppose you calculated a 90% confidence interval instead of the 95% confidence interval in part b. Would this new interval be wider or narrower? Why (be sure to include a mathematical explanation)?
   e. Suppose that researchers wanted to carry out a similar study of average caloric intake, but they want the width of a 90% confidence interval to be 80 calories wide. How many subjects should be sampled?

3. A large drug company has many potential new prescription drugs under clinical test. The company knows that 20% of all its drugs that reach this stage are eventually licensed for sale.
   a. Suppose the company plans to sample 100 of their new drugs, what is the probability that the proportion of drugs that are eventually licensed is at least 15%?
   b. Does it appear that the normal approximation for \( p \) is valid in this case? Explain.
   c. Would the probability in part a) be larger or smaller for a sample of size 250? Explain without actually calculating.
4. Forced expiratory volume (FEV) is an index of pulmonary function that measures the volume of air expelled after 1 second of constant effort. Assume that in 45-54 year-old nonsmoking men FEV is normally distributed with mean 4.0 liters and standard deviation of 0.7 liter. In comparably aged group of currently smoking men, FEV is normally distributed with a mean of 3.2 liters and a standard deviation of 0.8 liter. Suppose researchers are interested in a FEV of less than 2.5 liters because this level is regarded as showing some functional impairment (occasional breathlessness, inability to climb stairs, etc...)

a. What is the probability that a randomly selected currently smoking man has a functional impairment.

b. What percentage of non-smoking men have a functional impairment.

c. In a random sample of 5 smoking men, what is the probability that the average FEV is less than 2.5 liters.

d. Find the value of FEV in non-smoking men that represents the 67th percentile.

5. A building contractor has built a large number of houses of about the same size and value. The contractor claims that the average value of these houses is $265,000. A real estate appraiser randomly selects 35 of the new homes built by the contractor and assesses their average value as $263,500 with a standard deviation of $15,750.

a. Assuming normality, calculate a 90% confidence interval for the true average price for all new homes built by this contractor.

b. Does the confidence interval in part a) suggest that the contractor was wrong about the average value of these houses?

c. Test to see if the true mean value of these homes is significantly different from what the contractor claims. Test using $\alpha = 0.10$.

d. Suppose the data were not normally distributed. Does this effect the validity of the CI and hypothesis test? Explain.

6. Experts have predicted that 1 in 12 tractor-trailer units will be involved in an accident this year. One reason is that they speculate 1 in 3 tractor-trailer units has an imminently hazardous mechanical condition, related to the braking systems on the vehicle. A survey of 50 randomly selected tractor-trailer units passing through a weighing station confirmed that 19 had a potentially serious braking system problem.

a. Do the data provide evidence to show that mechanical braking problem has increased form what the experts believe? Test using $\alpha = 0.05$.

b. Do these data meet the assumptions required for the hypothesis test? Explain.

c. What would be worse a type I or type II error? Explain.

d. Suppose that the experts would like to further their researcher by examining the proportion of tractor-trailers that have hitching derailments. If they wanted to calculate a 99% confidence interval, with a bound on the error of no more than 5%, how many tractor-trailers should they sample?
IN ADDITION PLEASE STUDY THE FOLLOWING:

Your homework

Class notes

How to interpret Minitab output (normal probability plots, confidence intervals for the mean and hypothesis tests for the mean)

Topics:
  - population vs. sample vs. sampling distribution
  - Sampling distributions for the mean and proportion
  - parameters vs. statistics
  - appropriate notation
  - Central Limit Theorem (CLT)
  - Probabilities or percentages from Ch 7 and 8 using standardization
  - How to calculate confidence intervals
  - Appropriate conclusions for confidence intervals
  - Sample size calculations
  - One sample hypothesis tests for the mean and proportions, including all four parts
  - Type I and II errors
  - Appropriate assumptions for confidence intervals and hypothesis tests
  - How to use the z and t tables
ANSWER KEY

1. Suppose that the diameter of a population of Douglas fir trees in the Pacific Northwest is normally distributed with mean 50 cm and standard deviation 12 cm.

   a. What percentage of trees in the population have a diameter more than 75 cm?

   \[ X \geq 75 = Z \geq \frac{75 - 50}{12} = Z \geq 2.08 \]

   Look up 2.08 on Z table, 0.9812, so the percent more than 75 cm is 1 - 0.9812 = 0.0188

   b. A certain tree’s diameter is in the 33rd percentile. What is the diameter?

   Look inside the z table for 0.33. The z score associated with the 33rd percentile is -0.44, so now solve in terms of \( X \).

   \[-0.44 = \frac{X - 50}{12} \]

   \( X = 44.72 \) cm

   c. If 25 trees were randomly sampled from this population, what is the probability that their average diameter is less than 45 cm?

   \[ P \left( \overline{X} \leq 45 \right) = P \left( Z \leq \frac{45 - 50}{\sqrt{25}} \right) = P \left( Z \leq -2.08 \right) = 0.0188 \]

2. Two dietary instruments that are often used in studies to assess consumption of specific foods are the food-frequency questionnaire (FFQ) and the dietary record (DR). The major difference between the two methods is that the FFQ collects records of certain foods eaten over a certain period of time and then nutrient intake is estimated, while the DR calculates nutrient intake exactly from a record of every food ingested. The FFQ less expensive to administer but is considered less accurate than the DR (the gold standard of dietary analyses). To validate the FFQ, 173 nurses participated in a study to record their food intake using both methods and total caloric intake (per day) was obtained for the two methods. Descriptive statistics for each method can be found below (calor_dr for the DR and calor_ffq for the FFQ). Suppose that the data do not come from a normal distribution.

Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>TrMean</th>
<th>StDev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>calor_dr</td>
<td>173</td>
<td>1619.9</td>
<td>1606.0</td>
<td>1613.3</td>
<td>323.4</td>
<td>24.6</td>
</tr>
<tr>
<td>calr_ffq</td>
<td>173</td>
<td>1371.7</td>
<td>1297.6</td>
<td>1344.3</td>
<td>482.1</td>
<td>36.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>calor_dr</td>
<td>910.0</td>
<td>2518.0</td>
<td>1410.0</td>
<td>1784.5</td>
</tr>
<tr>
<td>calr_ffq</td>
<td>463.2</td>
<td>3077.3</td>
<td>1031.2</td>
<td>1591.0</td>
</tr>
</tbody>
</table>

   a. Calculate a 95% confidence interval for the true mean caloric intake for the FFQ method.
\[
\bar{x} \pm t \left( \frac{s}{\sqrt{n}} \right) = 1371.7 \pm 1.98 \left( \frac{482.1}{\sqrt{173}} \right) = 1371.7 \pm (1.98)(36.653) = (1299.24, 1444.16)
\]

b. It was mentioned above that the data do not come from a normal distribution, would you still consider the confidence interval that you calculated above valid? Why or why not?

Yes the interval above would still be considered valid because the sample size is large (n = 173). The central limit theorem guarantees that when n is large the sampling distribution of \( \bar{y} \) will be approximately normal regardless of the underlying distribution.

c. (TRUE or FALSE and say why) We are 95% confident that the average caloric intake of the 173 nurses using the FFQ method is between \( \text{interval calculated in part b} \).

FALSE, we know that the sample average caloric intake is between the interval calculated above.

d. Suppose you calculated a 90% confidence interval instead of the 95% confidence interval in part b. Would this new interval be wider or narrower? Why (be sure to include a mathematical explanation)?

A 90% confidence interval would be narrower than a 95% confidence interval because as the confidence goes down the interval becomes smaller. This is a result of the t multiplier becoming smaller with decreased confidence, so we would be adding and subtracting a smaller number to our estimate.

e. Suppose that researchers wanted to carry out a similar study of average caloric intake, but they want the width of a 90% confidence interval to be 80 calories wide. How many subjects should be sampled?

Using the sd from the current as a guess. Also if the confidence interval has a full width of 80, this means that the bound on the error is 40.

\[
n = \left( \frac{(1.645)(482.1)}{40} \right)^2 \\
= 393.08 \approx 394 \text{ Subjects}
\]

3. A large drug company has many potential new prescription drugs under clinical test. The company knows that 20% of all its drugs that reach this stage are eventually licensed for sale.

a. Suppose the company plans to sample 100 of their new drugs, what is the probability that the proportion of drugs that are eventually licensed is at least 15%?

\[
\mu_p = \pi = 0.20
\]

\[
\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.20(1-0.20)}{100}} = 0.04
\]

\[
P(p > 0.15) = P \left( z > \frac{0.15-0.20}{0.04} \right) = P(z > -1.25) = 1 - 0.1056 = 0.8944
\]

b. Does it appear that the normal approximation for \( p \) is valid in this case? Explain.

Yes because \( n(\pi) = 100(0.20) = 20 > 10 \) and \( n(1-\pi) = 100(0.80) = 80 > 10 \)

c. Would the probability in part a) be larger or smaller for a sample of size 250? Explain without actually calculating.
A larger sample size would cause the probability to be larger because the standard deviation would be smaller causing the z score to be further out in the tail. If the z score is pushed out to the tail this causes the probability to the left to be smaller and therefore 1 – this probability would be larger.

4. Forced expiratory volume (FEV) is an index of pulmonary function that measures the volume of air expelled after 1 second of constant effort. Assume that in 45-54 year-old nonsmoking men FEV is normally distributed with mean 4.0 liters and standard deviation of 0.7 liter. In comparably aged group of currently smoking men, FEV is normally distributed with a mean of 3.2 liters and a standard deviation of 0.8 liter. Suppose researchers are interested in a FEV of less than 2.5 liters because this level is regarded as showing some functional impairment (occasional breathlessness, inability to climb stairs, etc…)
   a. What is the probability that a randomly selected currently smoking man has a functional impairment.

\[
P(X < 2.5) = P\left(Z < \frac{2.5 - 3.2}{0.8}\right) = P(Z < -0.88) = 0.1894
\]

b. What percentage of non-smoking men have a functional impairment.

\[
X < 2.5 = Z < \frac{2.5 - 4}{0.7} = Z < -2.14 \quad 1.62\%
\]

c. In a random sample of 5 smoking men, what is the probability that the average FEV is less than 2.5 liters.

\[
P\left(\bar{X} < 2.5\right) = P\left(Z < \frac{2.5 - 3.2}{\frac{0.8}{\sqrt{5}}}\right) = P(Z < -1.96) = 0.025
\]

d. Find the value of FEV in non-smoking men that represents the 67th percentile.

The 67th percentile (0.67 below and 0.33 above) corresponds to a z score of 0.44

\[
0.44 = \frac{X - 4}{0.7} \quad X = 4.308 \text{Liters}
\]

5. A building contractor has built a large number of houses of about the same size and value. The contractor claims that the average value of these houses is $265,000. A real estate appraiser randomly selects 35 of the new homes built by the contractor and assesses their average value as $263,500 with a standard deviation of $15,750.
   a. Assuming normality, calculate a 90% confidence interval for the true average price for all new homes built by this contractor.

\[
\text{Df} = 35 - 1 = 34, \text{ use } 30 \\
263,500 \pm (1.70) \left(\frac{15750}{\sqrt{35}}\right) = (258974.20, 268025.80)
\]

b. Does the confidence interval in part a) suggest that the contractor was wrong about the average value of these houses?

Because the confidence interval contains his asserted value of $265,000 we cannot necessarily say that he is incorrect in his claim. The population mean value of all of these homes could very well be $265,000 because this
value is in the CI. However the true mean value could fall anywhere in this interval or not at all, but in the end all we can say is that the population mean does not appear to be significantly different from $265,000.

c. Test to see if the true mean value of these homes is significantly different from what the contractor claims. Test using $\alpha = 0.10$.

$$\text{Ho: } \mu = 265,000$$
$$\text{Ha: } \mu \neq 265,000$$

$$t_s = \frac{263500 - 265000}{15750/\sqrt{35}} = -0.56$$

Df = 35 - 1 = 34, use 30

$$2(0.310) < p < 2(0.277) = 0.620 < p < 0.554$$

Fail to reject Ho. These data do not provide sufficient evidence to suggest that there is a difference in the true mean value of these homes from $265,000 as the contractor claims.

d. Suppose the data were not normally distributed. Does this effect the validity of the CI and hypothesis test? Explain.

No, we do not need to assume that the population is normal when the sample size is larger than 30 because the central limit theorem will guarantee that the sampling distribution will be approximately normal.

6. Experts have predicted that 1 in 12 tractor-trailer units will be involved in an accident this year. One reason is that they speculate 1 in 3 tractor-trailer units has an imminently hazardous mechanical condition, related to the braking systems on the vehicle. A survey of 50 randomly selected tractor-trailer units passing through a weighing station confirmed that 19 had a potentially serious braking system problem.

a. Do the data provide evidence to show that mechanical braking problem has increased from what the experts believe? Test using $\alpha = 0.05$.

$$\text{Ho: } \pi = 1/3$$
$$\text{Ha: } \pi > 1/3$$

$$P = \frac{19}{50} = 0.38$$

$$z_s = \frac{0.38 - 0.33}{\sqrt{\left(\frac{0.33 \cdot 0.67}{50}\right)}} = \frac{0.05}{0.0665} = 0.751$$

$$p-value = 1-0.7734 = 0.2266 \quad \text{one tailed test, do not multiply p-value by 2!}$$

Fail to reject Ho. These data do not provide evidence to suggest that the proportion of tractor-trailers with mechanical braking problems have increased from what the experts believe of 33.3%.

b. Do these data meet the assumptions required for the hypothesis test? Explain.

1) Data are from a random sample:: Ok, stated in problem as “A survey of 50 randomly selected tractor-trailer units…”

2) Sample size is large:: $50(0.33) = 16.5 > 10$ and $50(1-0.33) = 33.5 > 10$, so ok
3) Sampling without replacement and sample size is no larger than 10% of the population:: In general they seem ok. It is doubtful that there was sampling with replacement (ie., the randomly sampled trucks were sampled more than once) and there are probably many trucks that drive through a weighing station, but we would want make sure that it is a station that supports enough tractor-trailer traffic to sample 50.

c. What would be worse a type I or type II error? Explain.

A type I error would be rejecting the null when the null is true. In the context of this problem this would be saying that the proportion of tractor-trailers that have braking problems is higher than experts believe, when it really isn’t higher.

A type II error would be failing to reject the null when the null is false. In the context of this problem this would be saying that the proportion of tractor-trailers that have braking problems has not increased from than experts believe, when they really have.

In this case it seems that a type II error would be much worse because we would think that there is not an increase in braking problems when there really is, and people’s lives would be put at risk. A type I error would lead us to believe there is a problem, and hopefully work to resolve the problem, when in reality the risk is not increased.

d. Suppose that the experts would like to further their researcher by examining the proportion of tractor-trailers that have hitching derailments. If they wanted to calculate a 99% confidence interval, with a bound on the error of no more than 5%, how many tractor-trailers should they sample?

Since we have no information about the proportion we will use $\pi = 0.5$

$$n = \pi(1-\pi)\left(\frac{z}{B}\right)^2 = 0.5(1-0.5)\left(\frac{2.58}{0.05}\right)^2 = 655.64 \approx 666 \text{ tractor-trailers}$$