Instructions: Answer each question to the best of your ability. Most questions require that you give a concluding or summary statement. These statements should be complete sentences. Good Luck!!! Have a Great Summer!

1. The following data (sorted for your convenience) on gross efficiency (ratio of work accomplished per minute to calorie expenditure per minute) for trained endurance cyclists was given in a paper in the journal *Medicine and Science in Sports and Exercise* (1992). Of interest to the researchers was to find a 95% confidence interval for $\mu$: mean gross efficiency for trained endurance cyclists.

```
GrossEff
18.3 18.9 19.0 19.9 20.1 20.1 20.5 20.5 20.5 20.5 20.6
20.8 20.9 21.2 21.4 21.9 22.1 22.6 22.6
```

a) What assumptions are needed to safely construct a 95% confidence interval in this case?

**THE SAMPLE NEEDS TO BE A SIMPLE RANDOM SAMPLE FROM THE NORMAL DISTRIBUTION**

b) Begin your assessment of these assumptions by finding the five number summary for this data.

<table>
<thead>
<tr>
<th>Lo</th>
<th>$Q_1$</th>
<th>Med</th>
<th>$Q_3$</th>
<th>Hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.3</td>
<td>20.1</td>
<td>20.5</td>
<td>21.4</td>
<td>22.6</td>
</tr>
</tbody>
</table>

c) Use the five number summary found above to construct a boxplot for this sample.

```
------------------
----------------------I    +         I-----------------
------------------
---+---------+---------+---------+---------+---------+----
GrossEff
18.40 19.20 20.00 20.80 21.60 22.40
```

**NOTE: THE BOXPLOT IS VERY SYMMETRIC LOOKING.**
d) Use the following Minitab output to continue in your assessment of the assumptions. Is it safe to conclude that our assumptions are reasonable?

**MTB > hist c1**
Histogram of GrossEff  N = 19

<table>
<thead>
<tr>
<th>Midpoint</th>
<th>Count</th>
<th>Leaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.5</td>
<td>1</td>
<td>18 3</td>
</tr>
<tr>
<td>19.0</td>
<td>2</td>
<td>18 9</td>
</tr>
<tr>
<td>19.5</td>
<td>0</td>
<td>19 0</td>
</tr>
<tr>
<td>20.0</td>
<td>3</td>
<td>19 9</td>
</tr>
<tr>
<td>20.5</td>
<td>5</td>
<td>20 11</td>
</tr>
<tr>
<td>21.0</td>
<td>3</td>
<td>20 5555689</td>
</tr>
<tr>
<td>21.5</td>
<td>1</td>
<td>21 24</td>
</tr>
<tr>
<td>22.0</td>
<td>2</td>
<td>21 9</td>
</tr>
<tr>
<td>22.5</td>
<td>2</td>
<td>22 1</td>
</tr>
</tbody>
</table>

**MTB > stem c1**
Stem-and-leaf of GrossEff  N = 19

<table>
<thead>
<tr>
<th>Leaf Unit = 0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>18 3</td>
</tr>
<tr>
<td>18 9</td>
</tr>
<tr>
<td>19 0</td>
</tr>
<tr>
<td>19 9</td>
</tr>
<tr>
<td>20 11</td>
</tr>
<tr>
<td>20 5555689</td>
</tr>
<tr>
<td>21 24</td>
</tr>
<tr>
<td>21 9</td>
</tr>
<tr>
<td>22 1</td>
</tr>
<tr>
<td>22 66</td>
</tr>
</tbody>
</table>

THE HISTOGRAMS AND STEMPLOTS INDICATE THAT THE DISTRIBUTION OF GROSSEFF'S ARE BASICALLY SYMMETRIC, NO OUTLIERS, BELL-SHAPED, UNIMODAL! IT IS SAFE TO ASSUME NORMALITY.

e) Verify that the interval that Minitab calculates is correct by building a 95% interval in this case.

\[
\bar{x} \pm t_{n-1}^{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 20.653 \pm 2.101 \cdot \frac{1.173}{\sqrt{19}} = (20.088, 21.218)
\]

WE ARE 95% CONFIDENT THAT THE MEAN GROSSEFF IS GREATER THAN 20.088 BUT NOT MORE THAN 21.218 FOR ENDURANCE CYCLISTS.

**MTB > tint 95 c1**

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
<th>95.0 % C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GrossEff</td>
<td>19</td>
<td>20.653</td>
<td>1.173</td>
<td>0.269</td>
<td>(20.087, 21.218)</td>
</tr>
</tbody>
</table>

f) Would it be safe to conclude, based on this interval, that the mean gross efficiency for all endurance cyclists is at least 20? Explain why or why not.

YES IT WOULD BE SAFE TO ASSUME. BOTH ENDPOINTS LIE HIGHER THAN THE VALUE 20
2. Many Americans think it doesn't matter which political party controls Congress. In an Associated Press article (1995) it was reported that 442 individuals in a sample of 1005 US adults said it wouldn't make much difference which party is in power. Find a 90% confidence interval for the proportion of US adults who believe that it wouldn't make much difference which party is in power by answering the following series of questions.

MTB > POne 1005 442;
SUBC>   Confidence 90;
SUBC>   UseZ.

**Test and CI for One Proportion**

Test of p = 0.5 vs p not = 0.5

<table>
<thead>
<tr>
<th>Sample</th>
<th>X</th>
<th>N</th>
<th>Sample p</th>
<th>90% CI</th>
<th>Z-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>442</td>
<td>1005</td>
<td>0.439801</td>
<td>(0.414047, 0.465555)</td>
<td>-3.82</td>
<td>0.000</td>
</tr>
</tbody>
</table>

b) What conclusion could you give about the proportion of US adults who believe that it wouldn't make much difference which party is in power, based on the 90% interval?

**WE ARE 90% CONFIDENT THAT THE TRUE PROPORTION OF PEOPLE WHO BELIEVE THAT IT DOESN'T MATTER WHICH PARTY CONTROLS THE CONGRESS IS AT LEAST .41 (41%) AND NOT MORE THAN .466 (46.6%).**

c) Based on your interval above, is it plausible that a majority of US adults feel that it makes no difference which party is in control? Explain your reasoning.

**NO IT IS NOT PLAUSIBLE, SINCE 0.5 IS NOT AN ELEMENT IN THE ABOVE INTERVAL.**
3. A new instructor at a university read an article from *The American Freshman* that discussed a study of the amount of time (in hours) college freshmen study each week. The study reported that the mean study time is 7.06 hours. The instructor feels that freshmen at her university study more than 7.06 hours per week on average and decides to test her hypothesis. She collects a simple random sample of size \( n=15 \) from her university and found the following data (in hours):

6.0, 1.7, 5.9, 13.3, 9.8, 7.8, 7.6, 10.3, 11.8, 11.0, 3.1, 12.3, 8.3, 10.4, 6.2

a) One of the assumptions to safely perform a test in this situation is for the data to come from a normal population. Since the sample size is so small, normality is often hard to determine conclusively. One of the most important departures from normality is the existence of many outliers. (The normal distribution is notorious for having few outliers, if any.) Based on the boxplot of the data below, do you believe it is safe to perform a test? Support your conclusion.

YES! THERE ARE NO OUTLIERS STARRED. THE DISTRIBUTION APPEARS TO BE RELATIVELY SYMMETRIC BASED ON EQUAL TAILS OF THE BOXPLOT AND THE MEDIAN IN THE MIDDLE OF THE BOX.

MTB > boxp c1

-------------------------
------------------------I   +   I-------------------------
-------------------------
------------------------+-----------------------------------------------+---Hours
  2.5    5.0    7.5   10.0   12.5

b) What are the appropriate null and alternative hypotheses for this problem?

\( H_0: \mu = 7.06 \quad vs \quad H_a: \mu > 7.06 \)

WHERE \( \mu \) IS THE MEAN HOURS PER WEEK "FRESHMEN" STUDY
MTB > ttest 7.06 c1;
SUBC> alte -1.
Test of mu = 7.060 vs mu < 7.060
Variable    N    Mean    StDev    SE Mean    T    P-Value
Hours       15   8.367   3.359     0.867     1.51       0.92

MTB > ttest 7.06 c1;
SUBC> alte 0.
Test of mu = 7.060 vs mu not = 7.060
Variable    N    Mean    StDev    SE Mean    T    P-Value
Hours       15   8.367   3.359     0.867     1.51       0.15

MTB > ttest 7.06 c1;
SUBC> alte 1.
Test of mu = 7.060 vs mu > 7.060
Variable    N    Mean    StDev    SE Mean    T    P-Value
Hours       15   8.367   3.359     0.867     1.51     0.077

---
c) Conduct the test mentioned in part (b), by using the Minitab output (circle the correct output) and by answering the following questions:
i) What is the test statistic? 1.51
ii) What is the p-value? 0.077
iii) What is my decision (reject $H_0$ or do not reject $H_0$)?

**SINCE 0.077 IS GREATER THAN 0.05, DO NOT REJECT $H_0$**

iv) What should the new instructor believe based on this research?

**THERE IS INSUFFICIENT EVIDENCE TO CONCLUDE THAT THE RESEARCHER IS CORRECT. THE MEAN TIME SPENT STUDYING AT HER UNIVERSITY IS NO MORE THAN OTHERS**

v) If the instructor had used as a level of significance $\alpha = 0.10$, would her conclusion be different? Explain.

**YES, BECAUSE 0.077 IS LESS THAN 0.1, WE WOULD REJECT $H_0$.**

vi) Verify that the test statistic calculated by Minitab is correct.

$$T = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{8.367 - 7.06}{\frac{3.359}{\sqrt{15}}} = \frac{1.307}{0.8673} = 1.51$$
4. In the past, it has been stated that 60 percent of all students at our college work part time (or more) during the school year. A counselor, after speaking with many students over the past few weeks regarding schedules, feels that this percentage is actually too low. We are interested in assessing if the proportion of students working part time is indeed larger than 0.60. A simple random sample of n = 120 students was obtained and of those 83 worked at least part time. Based on this sample can we safely conclude that the proportion is indeed larger than 0.60? Answer this general question by answering the following sequence of questions and by using the accompanying Minitab output.

**Seven Steps**

1. \( H_0: p = 0.6 \) vs \( H_a: p > 0.6 \)

2. **SIGNIFICANCE LEVEL:** CHOOSE SOMETHING THAT IS REASONABLE, SAY \( \alpha = 0.05 \)

3. Assumptions: A large SRS from a very large population.

   ASSUMPTIONS: SINCE THE POPULATION IS GREAT THAN 10*120 AND THE NUMBER OF SUCCESS =83 AND THE NUMBER OF FAILURES = 37 ARE BOTH GREATER THAN 10, THE ASSUMPTIONS ARE MET.

4. **TEST STATISTIC:** 
   
   \[
   Z = \frac{p - p_o}{\sqrt{p_o(1-p_o)/n}} = \frac{.69-.6}{\sqrt{.6(.4)/120}} = 2.05
   \]

5. **P-VALUE:** FROM MINITAB OUTPUT: P-VALUE = 0.021

6. **SINCE 0.021 IS LESS THAN \( \alpha = 0.05 \), REJECT \( H_0 \)**

7. **THERE IS STRONG EVIDENCE THAT THE TRUE PROPORTION OF STUDENTS WHO WORK AT LEAST PART TIME IS AT LEAST 0.60 (60%)**
MTB > POne 120 83;
SUBC> Test 0.6;
SUBC> Alternative -1;
SUBC> UseZ.

Test and CI for One Proportion

Test of p = 0.6 vs p < 0.6

<table>
<thead>
<tr>
<th>Sample</th>
<th>X</th>
<th>N</th>
<th>Sample p</th>
<th>Bound</th>
<th>Z-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83</td>
<td>120</td>
<td>0.691667</td>
<td>0.761008</td>
<td>2.05</td>
<td>0.980</td>
</tr>
</tbody>
</table>

MTB > POne 120 83;
SUBC> Test 0.6;
SUBC> UseZ.

Test and CI for One Proportion

Test of p = 0.6 vs p not = 0.6

<table>
<thead>
<tr>
<th>Sample</th>
<th>X</th>
<th>N</th>
<th>Sample p</th>
<th>95% CI</th>
<th>Z-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83</td>
<td>120</td>
<td>0.691667</td>
<td>(0.609041, 0.774293)</td>
<td>2.05</td>
<td>0.040</td>
</tr>
</tbody>
</table>

MTB > POne 120 83;
SUBC> Test 0.6;
SUBC> Alternative 1;
SUBC> UseZ.

Test and CI for One Proportion

Test of p = 0.6 vs p > 0.6

<table>
<thead>
<tr>
<th>Sample</th>
<th>X</th>
<th>N</th>
<th>Sample p</th>
<th>Bound</th>
<th>Z-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83</td>
<td>120</td>
<td>0.691667</td>
<td>0.622325</td>
<td>2.05</td>
<td>0.020</td>
</tr>
</tbody>
</table>

THE LAST TWO PROBLEMS DIDN'T APPLY TO OUR FINAL EXAM.