1. [16 pts] Short answer.
   a) [3] Statistics is about … _______________. (Fill in the blank with one word.)
   b) [2] A statistic is a numerical summary of a ________________.
      while a parameter is a numerical summary, of a ________________.
   c) [4] What does “90% confidence” mean?
   d) [2] For testing the hypothesis $H_0 : \mu = 50$ against the alternative $H_A : \mu > 50$, the value of the test statistic is $t = 2.131$ with df = 15. What is the P-value for the test?
   e) [2] Holding all other things the same, if the sample size is increased the width of the confidence interval will _________________.
   f) [3] Sketch a normal model with $\mu = 100$ and $\sigma = 15$. 
2. [24 pts] Multiple Choice:

a) ___ A 95% confidence interval for the mean distance walked per day is from 2.5 miles to 4.3 miles. The sample mean distance walked per day is?
   A: 2.5 miles  B: 3.4 miles  C: 3.0 miles  D: cannot be determined

b) ___ The correct interpretation of the interval in a) is?
   A: I am 95% confident that the sample mean is in the interval.
   B: 95% of the sample values are in the interval.
   C: 95% of the population values are in the interval.
   D: I am 95% confident that the population mean is in the interval.

c) ___ The P-value is …?
   A: The probability that the null hypothesis is rejected.
   B: The probability that the null hypothesis is true.
   C: The probability of observing a value of the test statistic more extreme than the one observed when the null hypothesis is true.
   D: The probability that the null hypothesis is not rejected.

d) ___ Holding all other things constant, as the level of confidence decreases, the width of a confidence interval …
   A: Stays the same
   B: Decreases.
   C: Increases.
   D: Cannot be determined.

e) ___ You have calculated the correlation coefficient between two variables to be –0.95. This would indicate?
   A: A strong linear relationship.
   B: A weak linear relationship.
   C: No relationship.
   D: No linear relationship.

f) ___ The time to complete this exam is an example of what type of variable?
   A: Categorical.
   B: Numerical – Continuous.
   C: Numerical – Discrete.


g) ___ Which of the following is not a measure of center?
   A: Mean.
   B: Median.
   C: Standard deviation.
   D: Midrange.

h) ___ Who was William Sealy Gosset?
   A: A famous statistician.
   B: A master brewer for Guinness.
   C: The person who developed Table T.
   D: All of the above.
3. [25 pts] The national average score for the ACT English subtest is 20.6 with a population standard deviation of $\sigma = 5$. Many high school students whose first language is not English take the ACT test. We wish to know if the population mean ACT English subtest score for students whose first language is not English is lower than the national average. A random sample of 40 students whose first language is not English has a sample mean ACT English subtest score of 18.3. ACT English subtest scores are normally distributed. Use $\alpha = 0.05$.

   a) [4] Set up the null and alternative hypothesis where $\mu$ is the population mean ACT English subtest score for students whose first language is not English.

   b) [4] State the test criteria for this problem.

   c) [5] Summarize the sample evidence by calculating the value of the test statistic.

   d) [4] Find the probability value.

   e) [4] Use the probability value to decide whether to reject or fail to reject the null hypothesis.

   f) [4] State a conclusion within the context of the problem.
4. [15 pts] In early November approximately 12,000 entertainment and television scriptwriters went on strike over their cut of the money from DVD sales and Internet sales. A FOX News/Opinion Dynamics Poll contacted 900 randomly selected registered voters on November 13-14, 2007 and asked them, “How disappointed are you that the writers strike will mean that you won't get new episodes of the television shows you usually watch?” 315 were disappointed, 450 were not disappointed and 135 either didn’t watch television or were unsure.


   c) [3] What proportion of the sample was disappointed?

   d) [6] Construct a 98% confidence interval for the population proportion of people who would be disappointed.

5. [5 pts] Consider a null hypothesis that it will rain today and an alternative that it will not rain today. Before leaving home you have to decide whether to take an umbrella or leave the umbrella at home. Fill in table below and indicate whether your decision is correct or an error.

   Hypothesis

<table>
<thead>
<tr>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
</tr>
<tr>
<td>e</td>
</tr>
<tr>
<td>c</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>s</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>o</td>
</tr>
<tr>
<td>n</td>
</tr>
</tbody>
</table>
6. [25 pts] The germination rate of a specific brand of soybeans is advertised to be greater than 92%. To test this claim, 250 soybean seeds are planted in a controlled environment and 235 of the 250 seeds germinate. Is this convincing evidence that the true population proportion of seeds that will germinate is greater than 92%? To answer this question, go through a step-by-step test of hypothesis. Use $\alpha = 0.05$.
   a) [4] Step 1: Set-up Hypotheses

   b) [4] Step 2: Test criteria

   c) [7] Step 3: Sample evidence

   d) [4] Step 4: Probability value

   e) [6] Step 5: Results

7. [5] For the following question identify the population, variable, type of variable and parameter. What is the average size of the high school graduating class for students attending Iowa State University?

   Population:

   Variable:

   Type of variable:

   Parameter:
8. [20 pts] Earlier in the semester we discussed an experiment looking at the efficacy of a drug prescribed to lower cholesterol. 36 individuals are involved in the experiment and the response is the change in their total cholesterol level. A negative value indicates that their cholesterol went down by that much. A positive value indicates that their cholesterol actually went up while taking the drug. Below is JMP output for the change in total cholesterol while taking the drug Lipator™.

\begin{itemize}
  \item a) [3] Describe the shape of the distribution of change in total cholesterol.
  \item b) [4] What is the value of the sample mean? the sample standard deviation?
  \item c) [6] Construct a 95% confidence interval on the population mean change in total cholesterol while taking the drug Lipator™.
\end{itemize}

e) [4] Based on the confidence interval in c), is the population mean change in total cholesterol while taking the drug Lipator™ zero? Explain briefly.

9. [15 pts] An article in the *Journal of the American Medical Association* examined whether the true mean body temperature is 98.6 degrees Fahrenheit. JMP was used to analyze the data on a random sample of 130 adults.

```
Quantiles
100.0% maximum 100.80 Std Dev 0.7331832
75.0% quartile 98.70 Std Err Mean 0.0643044
50.0% median 98.30 upper 95% Mean 98.376459
25.0% quartile 97.80 lower 95% Mean 98.122003
0.0% minimum 96.30 N 130

Test Mean=value
Hypothesized Value 98.6 Test Statistic -5.4548
Actual Estimate 98.2492 Prob > |t| <.0001
df 129 Prob > t 1.0000
Std Dev 0.73318 Prob < t <.0001
```
a) [3] Describe the shape of the distribution of body temperatures.

b) [4] Give the values of the sample mean and sample median body temperature. What does the comparison of these values tell you about the shape of the distribution? Explain briefly.

c) [8] Test the hypothesis that the population mean body temperature is 98.6° F against an alternative that the population mean is different from 98.6° F. Use the JMP output so that you do not have to do any additional calculations but be sure to give the steps for a test of hypothesis, e.g. null and alternative hypothesis, value of the test statistic, P-value, decision and reason for reaching that decision and a conclusion in the context of the problem. Use $\alpha = 0.05$. 