A person’s percentage body fat is determined from a person’s density. The density is obtained from the displacement of water in a large tub. In this exam we will look at men’s percentage body fat.

1. [27 pts] The American Council on Exercise (ACE) has a chart that describes different levels of percentage body fat. For male athletes the range of body fat is 6 to 13%. A random sample of 44 men has their percentage body fat determined by water displacement. The JMP analysis of the data is given below.

   a) [4] Looking at the histogram, describe the distribution of percentage body fat. Be sure to comment on shape, center and variability.
b) [3] Looking at the box plot, are their any potential outliers? How do you know this? If so, what is (are) the associated percentage body fat(s)?

c) [8] Could this sample be from a population of athletes? Test the hypothesis that the population mean percentage body fat is 13% versus an alternative that the population mean percentage fat is greater than 13%. Be sure to give the null and alternative hypothesis using appropriate statistical notation, value of the test statistic, P-value, decision and reason for reaching that decision and a conclusion in the context of the problem. For this problem use the usual significance level of 0.05.

d) [4] Give the values for the 95% confidence interval for the population mean percentage body fat. Explain briefly why this confidence interval is consistent with the test of hypothesis you did in c).
e) [4] Construct a 95% prediction interval for the percentage body fat of a randomly selected man. Note: the appropriate value of $t^*$ is 2.0167.

f) [4] What is the difference in interpretation between the confidence interval in d) and the prediction interval in e)?

2. [33 pts] The random sample of 44 men includes 24 men who are under 40 years of age and 20 men who are 40 to 55 years of age. The JMP analysis appears below.

![JMP analysis graph](image)

- Rsquare: 0.264098
- Adj Rsquare: 0.246576
- Root Mean Square Error: 8.087709
- Mean of Response: 18.77955
- Observations (or Sum Wgts): 44

<table>
<thead>
<tr>
<th>t Test</th>
<th>Assuming equal variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference</td>
<td>9.507</td>
</tr>
<tr>
<td>Std Err Dif</td>
<td>2.449</td>
</tr>
<tr>
<td>Upper CL Dif</td>
<td>14.448</td>
</tr>
<tr>
<td>Lower CL Dif</td>
<td>4.565</td>
</tr>
<tr>
<td>Confidence</td>
<td>0.95</td>
</tr>
</tbody>
</table>

- Level | Number | Mean | Std Dev | Std Err Mean | Lower 95% | Upper 95% |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>40 to 55</td>
<td>20</td>
<td>23.9650</td>
<td>10.2158</td>
<td>2.2843</td>
<td>19.184</td>
<td>28.746</td>
</tr>
<tr>
<td>under 40</td>
<td>24</td>
<td>14.4583</td>
<td>5.7648</td>
<td>1.1767</td>
<td>12.024</td>
<td>16.893</td>
</tr>
</tbody>
</table>

a) [5] Compare the percentage body fat of 40 to 55 year old men to that of the men under 40. Be sure to compare centers, variability and mention if there are any potential outliers in either group.
b) [3] What is the value of $s_p$, the pooled estimate of the common standard deviation, $\sigma$?

c) [8] Test the hypothesis that there is no difference between the population mean percentage body fat of 40 to 55 year old men and the population mean percentage body fat of men under 40 years old. Be sure to give the null and alternative hypothesis using clearly understood statistical symbols, value of the test statistic, $P$-value, decision and reason for reaching that decision and a conclusion in the context of the problem.

d) [5] Give the 95% confidence interval for the difference in population mean percentage body fat. What does this say about how much the population mean percentage body fat of men 40 to 55 years old differs from that of men under 40 years old?

e) [4] Is the condition of equal population standard deviations, $\sigma$, satisfied for these data? Support your answer.
Below is the JMP output for the distribution of the two-sample residuals.

f) [3] Looking at the normal quantile plot describe what you see and what this tells you about the condition that random errors are normally distributed.

g) [3] Looking at the box plot compare the median to the mean? What does this comparison indicate about the shape of the distribution of residuals?

h) [2] Looking at the histogram, where is the mound? How would you describe the shape of the distribution of residuals?
3. [40] Measuring percentage body fat by displacement of water is a time consuming process that requires the individual to be naked. Could a less time consuming and invasive measurement, like the circumference of a man’s abdomen (cm), be used to predict the percentage body fat? Below is JMP output looking at the relationship between abdomen circumference and percentage body fat.

![Graph showing relationship between Body Fat and Abdomen](image)

**Summary of Fit**
- RSquare: 0.706645
- RSquare Adj: 0.699660
- Root Mean Square Error: 5.10637
- Mean of Response: 18.77955
- Observations (or Sum Wgts): 44

**Parameter Estimates**

| Term               | Estimate | Std Error | t Ratio | Prob>|t| | Lower 95% | Upper 95% |
|--------------------|----------|-----------|---------|------|----------|-----------|
| Intercept          | –41.19   | 6.012     | –6.85   | <.0001* | –53.32   | –29.06    |
| Abdomen (cm)       | 0.648    | 0.0644    | 10.06   | <.0001* | 0.518    | 0.778     |

a) [3] Describe the general relationship between abdomen circumference and percentage body fat. Use complete sentences and say something about direction, form, strength and unusual values.

b) [3] Give the equation of the least squares regression line relating percentage body fat to the abdomen circumference.

c) [2] Use the equation in b) to predict the percentage body fat for a man with abdomen circumference of 110 cm.
d) [5] Calculate a 95% prediction interval for a man with abdomen circumference of 110 cm. Note $t^* = 2.01081$. Note: the sample mean abdomen circumference is 92.6 cm and the sample variance of abdomen circumference is 146.21 cm$^2$.

e) [5] Give an interpretation of the estimated slope coefficient within the context of the problem.

f) [3] Why doesn’t the intercept have an interpretation within the context of the problem?

g) [3] Give the value of $R^2$ and an interpretation of that value within the context of the problem.

h) [2] Give the value of the estimate of the random error standard deviation, $\sigma$. 
i) [6] Report the 95% confidence interval for the slope. Use this interval to test for a statistically significant relationship between percentage body fat and abdomen circumference.

j) [4] Describe what you see in the plot of residuals versus predicted body fat. What does this plot tell you about the adequacy of the linear model?

![Residuals Versus Predicted Body Fat Plot]

k) [4] Comment on the condition of normally distributed random errors. Be sure to support your comments by referring to the normal quantile plot of residual.

![Normal Quantile Plot]