Exam I

You have one hour for this exam. The exam is worth 15 percent of your grade. The exam is closed book, but calculators are permitted. Please do all your work carefully and check your calculations before handing your work in. PLEASE SHOW ALL YOUR WORK TO AID IN THE AWARDING OF PARTIAL CREDIT WHERE APPLICABLE.
1. (20 points) Data were gathered on the change in unemployment and the change in Real GDP from 1998-2004. The regression was run and the results are presented below.

**Regression Analysis: ChUnem versus %chGDP**

The regression equation is

\[
\text{ChUnem} = 1.23 - 0.411 \times \text{%chGDP}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.2301</td>
<td>0.1501</td>
<td>8.19</td>
</tr>
<tr>
<td>%chGDP</td>
<td>-0.41075</td>
<td>0.04237</td>
<td>-9.69</td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>14.386</td>
<td>14.386</td>
<td>93.98</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>23</td>
<td>3.521</td>
<td>0.153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>17.906</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Whoops, I accidentally cut out the $R^2$. From the output above, please calculate the $R^2$.

\[
R^2 = \frac{\text{sum of sq. Res}}{\text{sum of sq. Total}} = \frac{14.386}{17.906} = 80.32
\]

b) In terms of the variables, please explain what the $R^2$ tells us.

80.32% of the variation in the change in unemployment is explained by the change in GDP.

c) At a .05 level please test whether the slope is $> 0$. Show your work

\[
\begin{align*}
H_0: \beta_2 & \leq 0 \\
\text{vs} \\
H_a: \beta_2 & > 0
\end{align*}
\]

\[
\text{Reject } H_0 \text{ if } (t > t_{n-2, .05} = 1.711) \Rightarrow \beta_2 \geq t > 1.711
\]

\[
-0.41075 
\]

Do not reject $H_0$. 

2. (44 points) Once again, neither the Cleveland Browns nor the Detroit Lions made it to the Superbowl. I gathered data for all 32 teams for the 2005 season on

\[
\begin{align*}
\text{Wins} & = \text{Number of Games Won} \\
\text{PtsFor} & = \text{Points Scored by the Team} \\
\text{PtsAgst} & = \text{Points Scored against the Team}
\end{align*}
\]

I ran a regression and found,

\textbf{Regression Analysis: Wins versus PtsFor, PtsAgst}

The regression equation is

\[
\text{Wins} = 10.3 + 0.0272 \text{PtsFor} - 0.0340 \text{PtsAgst}
\]

\[
\begin{array}{lcccc}
\text{Predictor} & \text{Coef} & \text{SE Coef} & T & P \\
\text{Constant} & 10.267 & 1.932 & 5.31 & 0.000 \\
\text{PtsFor} & 0.027159 & 0.003257 & 8.34 & 0.000 \\
\text{PtsAgst} & -0.034031 & 0.003673 & -9.27 & 0.000 \\
\end{array}
\]

\[
S = 1.10280 \quad \text{R-Sq} = 90.1\% \quad \text{R-Sq(adj)} = 89.4\%
\]

\textbf{Analysis of Variance}

\[
\begin{array}{lcccc}
\text{Source} & \text{DF} & \text{SS} & \text{MS} & \text{F} & \text{P} \\
\text{Regression} & 2 & 320.73 & 160.37 & 131.86 & 0.000 \\
\text{Residual Error} & 29 & 35.27 & 1.22 & & \\
\text{Total} & 31 & 356.00 & & & \\
\end{array}
\]

\[
\text{Source} \quad \text{DF} \quad \text{Seq SS} \\
\text{PtsFor} & 1 & 216.30 \\
\text{PtsAgst} & 1 & 104.43 \\
\]

a) In terms of the variables, please explain what the slope on PtsFor tells us.

\[
\text{b) The Cleveland Browns won} \, g \, \text{games this past season. They scored 232 points and gave up 301. How many games did the model predict they would win?}
\]

\[
\begin{align*}
\text{Wins} &= 10.3 + 0.0272(232) - 0.0340(301) \\
&= 10.3 + 6.31 - 10.274 \\
&= 6.376
\end{align*}
\]

\[
\text{c) What is the error, i.e., the residual for the Browns. (No points for saying the error was that their fans thought they would do better.)}
\]

\[
\begin{align*}
\text{Residual} &= Y - \hat{Y} = 6 - 6.376 = -0.376
\end{align*}
\]

(This problem continues on the next page.)
d) The Browns scored the fewest points of any team in the league; 232. (Ouch!) That was 108 points less than the league average. How many additional games would they have been predicted to win, compared to your answer in part b), had they scored the league average number of points?

\[108 \div 20 = 5.4\]

\[x = 5.4\]

\[\text{additional games won}\]

e) Suppose that the Coach claims that offense doesn’t win games, defense does. He goes on to state that the number of points scored does not matter. Based on the formal model above, test the idea that the number of points scored does not affect the number of games won. Please use a .05 level, .025 under each tail.

\[H_0: \beta_2 = 0 \quad H_1: \beta_2 < 0\]

\[b = \frac{-32}{1.2} = -26.67\]

\[t = \frac{-26.67}{3.25} = -8.2\]

\[t < -2.045\]

\[\text{do not reject } H_0\]

f) In terms of the variables, please explain what your results in part e) mean.

There is less than a 1% chance that points scored has no effect on games won.

g) Suppose your roommate says that points scored and allowed don’t matter, all that matters is wins. You claim that these two variables do in fact matter as they determine wins. Your roommate still says they don’t matter. Formally, test the idea at the .05 level that

\[H_0: B_2 = B_3 = 0 \quad \text{vs.} \quad H_A: H_0 \text{ is false}\]

\[F = \frac{5.54}{2.9} = 1.89\]

\[\text{do not reject } H_0\]

\[\text{at } 0.05 \text{ level}\]
3-5 are worth 12 points each.

3.a) Suppose I wanted to find the mean of a population. What would be an example of an unbiased estimator from a sample? (Formula of one, please.)

\[
\bar{X} = \frac{\sum X_i}{n}
\]

The sample mean is an unbiased estimator (also median).

b) What would be an example of a biased one?

\[
E(X_i + 1) = \frac{\sum (X_i + 1)}{n}
\]

On anything with a cat, u is the population mean.

c) What does it mean for an estimator to be biased?

\[
E(\text{estimator}) = E(\bar{X}) \neq \mu
\]

4. Suppose we look at the probability distribution of males and females that were admitted to an Ivy League graduate school. There are two programs; Arts and Sciences.

<table>
<thead>
<tr>
<th>Program</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>.06</td>
<td>.20</td>
</tr>
<tr>
<td>Sciences</td>
<td>.51</td>
<td>.23</td>
</tr>
</tbody>
</table>

a) What does the figure .20 mean in terms of the variables?

20% of the admitted students were females.

b) Calculate the marginal probabilities.

c) Are gender and program independent? Demonstrate.

\[
\text{if independent } P(\text{Male \& Arts}) = P(\text{Male}) \cdot P(\text{Arts})
\]

.06 \neq .57 \cdot .26

.06 \neq .1482

\[\Rightarrow\text{ not independent}\]
5. Check your OWU e-mail. There should be a file that has been mailed to you as an Excel file. Cut and paste the variables into Minitab. The variables are:

**DEMAND FOR CHICKENS, UNITED STATES, 1960-1982**

\( Y \) = Per Capita Consumption of Chickens, Pounds  
\( X_2 \) = Real Disposable Income Per Capita, $  
\( X_3 \) = Real Retail Price of Chicken Per Pound, Cents  
\( X_4 \) = Real Retail Price of Pork Per Pound, Cents  
\( X_5 \) = Real Retail Price of Beef Per Pound, Cents  
\( X_6 \) = Composite Real Price of Chicken Substitutes Per Pound, Cents

a) Run a regression of \( Y \) on all of the \( X \)'s. Write the regression equation below.

\[
y = 38.6 + 0.8014 Y_2 - 0.652 X_3 + 0.243 X_4 + 0.104 X_5 - 0.071 X_6
\]

\( (0.99) \quad (0.79) \quad (2.78) \quad (1.45) \quad (-0.78) \)

b) Please run a test to see if we can drop all of the those variables (together) that are not statistically significant at a .05 level. Show your work and any relevant test statistics.

\[
\bar{F} \text{ indp vars } k-i = 5 \quad k=6
\]

\[
H_0: \beta_2 = 0 \quad \text{ vs } \quad H_a: \beta_2 \neq 0
\]

\[
H_0: \beta_3 = \beta_5 = \beta_6 = 0 \quad \text{ vs } \quad H_a: H_0 \text{ is false}
\]

\[
F_{k-i, n-k-i-1} = F_{5,17,0.05} = 2.20
\]

If \( F > 2.20 \) reject \( H_0 \)

\[
\frac{SSR_{H_0}}{I-k-1} = \frac{1129.31}{6} = 188.22
\]

\[
\frac{SSR_{H_0}}{n-k} = \frac{66.62}{17} = 3.91 > 2.20
\]

reject \( H_0 \)