Econ 371  
Exam #2 - Practice

**Multiple Choice (5 points each):** For each of the following, select the single most appropriate option to complete the statement.

1) The reason why estimators have a sampling distribution is that  
   a) economics is not a precise science.  
   b) individuals respond differently to incentives.  
   c) in real life you typically get to sample many times.  
   d) the values of the explanatory variable and the error term differ across samples.

2) Multiplying the dependent variable by 100 and the explanatory variable by 100,000 leaves the  
   a) OLS estimate of the slope the same.  
   b) OLS estimate of the intercept the same.  
   c) regression R2 the same.  
   d) variance of the OLS estimators the same.

3) Consider the following regression line: \( \text{TestScore} = 698.9 - 2.28 \times STR \). You are told that the t-statistic on the slope coefficient is 4.38. What is the standard error of the slope coefficient?  
   a) 0.52  
   b) 1.96  
   c) -1.96  
   d) 4.38

4) The confidence interval for the sample regression function slope  
   a) can be used to conduct a test about a hypothesized population regression function slope.  
   b) can be used to compare the value of the slope relative to that of the intercept.  
   c) adds and subtracts 1.96 from the slope.  
   d) allows you to make statements about the economic importance of your estimate.

5) Under imperfect multicollinearity  
   a) the OLS estimator cannot be computed.  
   b) two or more of the regressors are highly correlated.  
   c) the OLS estimator is biased even in samples of n > 100.  
   d) the error terms are highly, but not perfectly, correlated.

6) In the multiple regression model, the adjusted \( R^2, \overline{R^2} \)  
   a) cannot be negative.  
   b) will never be greater than the regression \( R^2 \).  
   c) equals the square of the correlation coefficient r.  
   d) cannot decrease when an additional explanatory variable is added.
7) When testing joint hypothesis, you should
a) use t-statistics for each hypothesis and reject the null hypothesis if all of the restrictions fail.
b) use the F-statistic and reject all the hypothesis if the statistic exceeds the critical value.
c) use t-statistics for each hypothesis and reject the null hypothesis once the statistic exceeds the critical value for a single hypothesis.
d) use the F-statistics and reject at least one of the hypothesis if the statistic exceeds the critical value.

8) When testing the null hypothesis that two regression slopes are zero simultaneously, then you cannot reject the null hypothesis at the 5% level, if the ellipse contains the point
a) (-1.96, 1.96).
b) |(0, 1.96)|.
c) (0,0).
d) (1.962, 1.962).

Problems: Provide the requested information for each of the following questions. Be sure to show your work.

9) (continuation from Chapter 4, number 3) You have obtained a sub-sample of 1744 individuals from the Current Population Survey (CPS) and are interested in the relationship between weekly earnings and age. The regression, using heteroskedasticity-robust standard errors, yielded the following result:

\[ \hat{Earn} = 239.16 + 5.20 \times Age, \quad R^2 = 0.05, \quad SER = 287.21, \]
\[ (20.24) \quad (0.57) \]

where \( Earn \) and \( Age \) are measured in dollars and years respectively.

(a) Interpret the results.

(b) Is the effect of age on earnings large?

(c) Is the relationship between Age and Earn statistically significant?

(d) Construct a 95% confidence interval for both the slope and the intercept.
9) Attendance at sports events depends on various factors. Teams typically do not change ticket
prices from game to game to attract more spectators to less attractive games. However, there
are other marketing tools used, such as fireworks, free hats, etc., for this purpose. You work
as a consultant for a sports team, the Los Angeles Dodgers, to help them forecast attendance,
so that they can potentially devise strategies for price discrimination. After collecting data
over two years for every one of the 162 home games of the 2000 and 2001 season, you run
the following regression:

\[
\text{Attend} = 15,005 + 201 \times \text{Temperat} + 465 \times \text{DodgNetWin} + 82 \times \text{OppNetWin} \\
+ 9647 \times \text{DFSaSu} + 1328 \times \text{Drain} + 1609 \times D150m + 271 \times DDiv - 978 \times D2001; \\
(8,770) \quad (121) \quad (169) \quad (26) \quad (1505) \quad (3355) \quad (1819) \quad (1,184) \quad (1,143)
\]

\[R^2=0.416, \text{SER} = 6983\]

where \(\text{Attend}\) is announced stadium attendance, \(\text{Temperat}\) it the average temperature on
game day, \(\text{DodgNetWin}\) are the net wins of the Dodgers before the game (wins-losses),
\(\text{OppNetWin}\) is the opposing team’s net wins at the end of the previous season, and
\(\text{DFSaSu}, \text{Drain}, D150m, Ddiv, \text{and} D2001\) are binary variables, taking a value of 1 if the
game was played on a weekend, it rained during that day, the opposing team was within a
150 mile radius, the opposing team plays in the same division as the Dodgers, and the
game was played during 2001, respectively. Numbers in parentheses are
heteroskedasticity- robust standard errors.

(a) Interpret the regression results. Do the coefficients have the expected signs?

(b) Are the slope coefficients statistically significant?

(c) To test whether the effect of the last four binary variables is significant, you have your
regression program calculate the relevant \(F\)-statistic, which is 0.295. What is the critical
value? What is your decision about excluding these variables?

(d) Excluding the last four binary variables results in the following regression result:

\[
\text{Attend} = 14,838 + 202 \times \text{Temperat} + 435 \times \text{DodgNetWin} + 90 \times \text{OppNetWin} \\
+ 10,472 \times \text{DFSaSu}, \quad R^2=0.410, \text{SER} = 6925
\]
According to this regression, what is your forecast of the change in attendance if the temperature increases by 30 degrees? Is it likely that people attend more games if the temperature increases? Is it possible that Temperat picks up the effect of an omitted variable?

(e) Assuming that ticket sales depend on prices, what would your policy advice be for the Dodgers to increase attendance?