Instructor: Prof. Murdock

Duration: 50 minutes: 45 minutes are for writing the quiz and 5 minutes are for the TA's and me to distribute and collect papers while you are silently seated with writing instruments down on your desk.

Format: 15 multiple-choice questions. Record answers on pink SCANTRON form. Correct answers worth 5 points each and incorrect answers worth 0 points. 75 total possible points.

Allowed aids: A non-programmable calculator and attached aid sheets.

Instructions:

- Answers must be properly recorded on the pink SCANTRON form to earn marks
- Print your **LAST NAME** and **INITIALS** in the boxes AND darken each letter in the corresponding bracket below each box; Sign your name in the **SIGNATURE** box
- Print your 9 digit **STUDENT NUMBER** in the boxes AND darken each number in the corresponding bracket below each box
- Your **FORM NUMBER** is **01**
- Use only a pencil or blue or black ball point pen
    - **Pencil strongly recommended** because it can be erased
- Make dark solid marks that fill the bubble completely
- Erase completely any marks you want to change
- Crossing out a marked box is not acceptable and is incorrect
- If more than one answer is selected then that question earns 0 points

****Unless otherwise specified use the conventional 5 percent significance level.****

(1) A linear regression model makes which of these assumptions?

Assumption 1: The variance of the error is constant across observations

Assumption 2: Only linear transformations have been applied to the variables

Assumption 3: There is no correlation between the error term and the dependent variable

(A) Only 1
(B) Only 2
(C) Only 3
(D) Only 1 and 3
(E) All three: 1, 2 and 3

► **Questions (2) – (4):** This multiple regression model is estimated with a sample size of 900.

$$y_{it} = \alpha + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \varepsilon_{it}$$

(2) To infer that the model is statistically significant overall the *F* test statistic must be _____.

(A) smaller than 2.1
(B) greater than 2.1
(C) greater than 2.6
(D) smaller than 3.0
(E) greater than 3.0

(3) To infer that β_2 is less than 10 the *t* test statistic must be smaller than _____.

(A) -2.353
(B) -1.960
(C) -1.645
(D) 7.897
(E) 8.014

(4) You can recognize that the data are _____.

(A) panel
(B) time series
(C) experimental
(D) observational
(E) cross sectional

(5) For a simple regression a low value of the coefficient of determination (R^2) suggests that _____.

(A) the SSE may be smaller than the SSR
(B) the standard error of estimate may be low
(C) there may be a strong positive correlation
(D) there may be a strong negative correlation
(E) the model may not be statistically significant

► **Question (6):** A researcher estimates a multiple regression model using cross sectional data on households to describe the relationship between the amount of household garbage in an average week (pounds), the size of the house (square feet), number of children, and number of adults. Here is an excerpt from the Excel analysis. The model overall is statistically significant.

16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	
17	Intercept	7.19	1.09	6.59	1.30E-10	
18	HSize	0.0019	0.0006	3.21	0.0014	
19	Children	1.10	0.14	7.84	3.58E-14	
20	Adults	1.04	0.23	4.48	9.58E-06	

(6) Which is the *best* interpretation of the coefficient 0.0019 on the house size variable?

(A) A house that is 1000 square feet will produce 1.9 pounds of garbage on average

(B) Increasing the size of a house by 1000 square feet will increase garbage by 1.9 pounds per week

(C) Increasing house size by 1000 square feet holding all observed and unobserved variables constant means 1.9 pounds more of garbage per week on average

(D) Houses that are 1000 square feet bigger typically produce 1.9 more pounds of garbage per week even though we have controlled for the number of children and adults

(E) Houses that are 1000 square feet larger tend to have more children and adults and therefore they produce more garbage: on average an extra 1.9 pounds per week

► **Questions (7) – (9):** Consider this STATA output. Parts have been intentionally erased.

```
      Source |       SS       df       MS              Number of obs =     125
-------------+------------------------------           F(  4,   120) =    2.28
       Model |   1224.5546     4   306.13865
    Residual |   16139.966   120  134.499717
-------------+------------------------------
       Total |  17364.5206   124  140.036457

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.459029   .9809314     1.49   0.140    -.4831465    3.401205
          x2 |   2.618772   1.105818     2.37   0.019     .4293292    4.808214
          x3 |  -1.117076    1.04493    -1.07   0.287    -3.185965    .9518135
          x4 |  -.3368922   1.060085    -0.32   0.751    -2.435786    1.762002
       _cons |   24.18203   1.063708    22.73   0.000     22.07596    26.28809
------------------------------------------------------------------------------
```

(7) What are the critical values for the test of statistical significance for β_1?

(A) -1.490 and 1.490
(B) -1.645 and 1.645
(C) -1.658 and 1.658
(D) -1.960 and 1.960
(E) -1.980 and 1.980

(8) What is the test statistic for the hypothesis test H_0: $\beta_2 = 1$ and H_1: $\beta_2 > 1$?

(A) 1.37
(B) 1.46
(C) 1.49
(D) 1.96
(E) 2.37

(9) What is the approximate p-value for the overall test of statistical significance of the model?

(A) ≈ 0
(B) a bit less than 0.01
(C) between 0.01 and 0.025
(D) between 0.025 and 0.05
(E) between 0.05 and 0.10

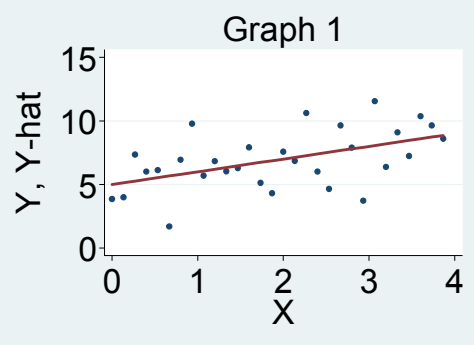
(10) To encourage students a professor sends a message to some students after the first test.

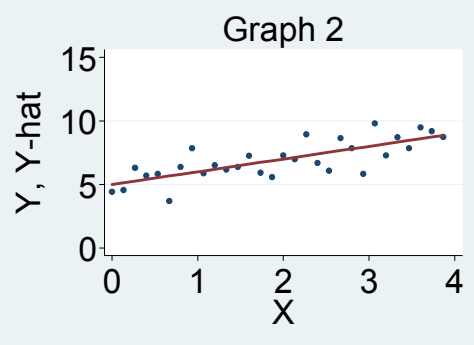
- If a student did very badly the message gives advice for making study time more effective.
- If a student did very well the message is congratulatory.

On the next test students that had received the study advice message on average did better. However, students that had received the congratulatory message on average did worse. The results are statistically significant. What should the professor conclude?

(A) The study advice message is effective
(B) The study advice message is not effective
(C) Stop sending congratulatory messages because they are harmful
(D) These messages may not have any effect on student performance
(E) The congratulatory message probably leads students to work less hard for the next test

► **Questions (11):** Consider these two scatter diagrams and OLS lines.





(11) Compared to Graph 1, Graph 2 clearly shows a case where the standard error of estimate is _____ and the standard error of the slope is _____.

(A) similar; similar
(B) similar; smaller
(C) bigger; smaller
(D) smaller; similar
(E) smaller; smaller

► **Questions (12) – (15):** In experimental data the dependent variable measures the number of coughs per hour. The independent variable is the number of milliliters (mL) of a cough syrup.

X

	Percentiles	Smallest		
1%	0	0		
5%	2.5	0		
10%	7.5	5	Obs	40
25%	15	5	Sum of Wgt.	40
50%	20		Mean	21.5
		Largest	Std. Dev.	11.83216
75%	32.5	40		
90%	40	40	Variance	140
95%	40	40	Skewness	.1065198
99%	40	40	Kurtosis	1.9443

Y

	Percentiles	Smallest		
1%	1	1		
5%	2	2		
10%	2.5	2	Obs	40
25%	3.5	2	Sum of Wgt.	40
50%	5		Mean	5
		Largest	Std. Dev.	2.12434
75%	6	9		
90%	8.5	9	Variance	4.512821
95%	9	9	Skewness	.4875661
99%	10	10	Kurtosis	2.768595

$$\underset{(0.6600)\quad(0.0270)}{\textit{Y-hat} = 6.4570 - 0.0678^{*}X}$$

standard error of estimate: 1.9929

(12) Is there sufficient evidence to conclude that increasing cough syrup by 1 mL decreases the number of coughs by at least 0.02 per hour (i.e. to infer that the research hypothesis H_1: $\beta < -0.02$ is true)?

(A) No, not with any significance level ≤ 10%
(B) Yes, with a significance level ≤ 1%
(C) Yes, with a significance level > 1% and ≤ 2.5%
(D) Yes, with a significance level > 2.5% and ≤ 5%
(E) Yes, with a significance level > 5% and ≤ 10%

(13) We can be 98% confident that an extra mL of cough syrup reduces the number of coughs per hour by _____ coughs.

(A) 0.002 to 0.133
(B) 0.006 to 0.141
(C) 1.247 to 9.582
(D) 4.219 to 5.781
(E) 4.775 to 4.911

(14) If you give a person 30 mL of cough syrup you can be 90% confident that that person will cough between _____ times per hour.

(A) 1.0 to 7.8
(B) 2.0 to 9.0
(C) 2.2 to 6.5
(D) 2.7 to 6.1
(E) 3.4 to 5.4

(15) Suppose someone has a psychological disorder and coughs 100 times per hour. If that person had been randomly sampled and randomly assigned to receive zero mL of cough syrup how would this affect our analysis?

Effect 1: We would overestimate the effectiveness of the drug

Effect 2: We would underestimate the effectiveness of the drug

Effect 3: We would violate the assumption that the error has a mean of zero

(A) Only 1
(B) Only 2
(C) Only 3
(D) Only 1 and 3
(E) Only 2 and 3

Your form number is **01**. Complete the FORM box at the top right of your pink SCANTRON form.

BLANK PAGE