

READ THIS SECTION CAREFULLY.

This exam is for your work ONLY. You may use:

1. Your notes
2. Your book
3. The class videos (made by me)
4. A calculator

You may not use:

1. Your computer (other than downloading, printing, or reading the exam). Note: If you really want, you can type your answers in Word, but I don't really think you will want to with all of the equations.
2. Excel or any other spreadsheet program
3. The internet except D2L notes and videos
4. Other individuals

Failure to follow these guidelines will result in:

1. A grade of zero on the exam
2. A report of academic misconduct to the Office of Student Standards

Please believe me when I say that I take this seriously. I need to be able to trust that this is your work. Just because this is an online class does not change anything – you should treat this as you would an in-class exam where you are allowed to use your book and notes.

Completion of this exam means you have acknowledged the rules listed above and will act accordingly.

****You will upload your answers either in Word or a pdf the same way as you submit your homework assignments in ONE document.** NO PICTURES! Make sure it is in order!!**

Due date: Tuesday, September 22, 2020 at 11:00 pm. No late exams will be accepted!

Please make sure your answers are written clearly and IN ORDER. Make sure every question is labeled with the proper number. Read each question carefully. For each hypothesis test, you must write both the null and alternative hypotheses, the decision rule, the calculated test statistic, the critical value of the test statistic and then (finally) the decision. You must have ALL parts! (Don't assume that I will find it – make it clear.)

Show all work if you want any partial credit. Point values are in parentheses.

1. You are hired as a consultant to the Bob's Bargain Bin to determine the factors that affect sales revenues at each of their stores. You estimate a model using OLS with a cross-section data set containing 46 observations (their current stores). The estimated model was:

$$Sales_i = \beta_0 + \beta_1 Pop_i + \beta_2 Income_i + \beta_3 Sunday_i + \beta_4 Compdist_i + \varepsilon_i$$

where:

$Sales_i$ = revenue for store i (thousands of \$)

Pop_i = population in city where store i is located (in thousands)

$Income_i$ = median income of population in city where store i is located (thousands of \$)

$Sunday_i$ = 1 if store i is open on Sunday; zero otherwise

$Compdist_i$ = miles to closest competitor for store i

The estimated results are:

<u>Coefficients:</u>		
	<u>Estimate</u>	<u>Std. Error</u>
(Intercept)	34015.78	12770.09
Pop	0.9782	0.41998
Income	0.64981	0.44298
Sunday	2498.67	1123.86
Compdist	434.978	135.987

R-squared: 0.4895, Adjusted R-squared: 0.4653		
F-statistic: 19.93 on 4 and 41 DF		

(6) a. Test the overall significance of the model at the 5% significance level.

(4) b. Provide an interpretation of the coefficient on $Sunday_i$.

(6) c. It is difficult to predict the sign on the coefficient for $Income_i$. After all, this is a bargain store. While shopping may be a normal good, it is also possible that as income rises, fewer people would want to shop at this particular store. Therefore, we do not have a hypothesis for the sign of the coefficient for $Income_i$. However, we would like to test if the coefficient is zero. Do so (at the 5% significance level).

(6) d. At the same time, we believe that the distance to the nearest competitor's store will positively impact the sales at Bob's Bargain Bin. Test this hypothesis (at the 5% level).

(4) e. The R^2 of the model estimated is 0.4895. What does this number tell us? Be specific.

2. Suppose you are interested in determining the factors that affect the box office revenues for movies produced in 2018. You gather information on 85 movies and estimate the following model:

$$Rev_i = \beta_0 + \beta_1 Budget_i + \beta_2 Rating_i + \beta_3 Action_i + \beta_4 Summer_i + \varepsilon_i$$

where:

Rev_i = box-office revenues for movie i (in thousands of dollars)

$Budget_i$ = production budget of film i (in thousands of dollars)

$Rating_i$ = 1 if movie i is rated R; = 0 otherwise

$Action_i$ = 1 if movie i is an action movie; = 0 otherwise

$Summer_i$ = 1 if movie i was released during the summer; = 0 otherwise

The results from the estimation are as follows:

<u>Coefficients:</u>		
	<u>Estimate</u>	<u>Std. Error</u>
(Intercept)	12.322	2.566
Budget	2.589	1.645
Rating	-52.887	19.582
Action	12.876	5.998
Summer	37.921	16.776

R-squared: 0.7423, Adjusted R-squared: 0.7183		
F-statistic: 21.62 on 4 and 80 DF		

(6) a. It is believed that the impact of budget on box-office revenue is positive. Test this hypothesis at the 5% level.

(4) b. Interpret the coefficient on $Action_i$. (Be careful of units!)

(6) c. Movies that have R ratings generally appeal to fewer customers. Test this hypothesis at the 5% level.

(6) d. Having a movie release during the summertime is expected to raise box-office revenues by more than \$30,000. Test this hypothesis at the 5% level.

3. You are a macroeconomist working for the government. You have been assigned the task to estimate a model to better understand the factors that affect consumption over time. You gather data from the past 50 years and estimate the following model:

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 T_t + \beta_3 Conf_t + \varepsilon_t$$

where:

$Const_t$ = annual consumption expenditure in time t (in millions of dollars)

Y_t = real GDP for time t (in millions of dollars)

T_t = annual taxes paid in year t (in millions of dollars)

$Conf_t$ = consumer confidence in economy in time t (scale of 0 to 100)

The results from the estimation are as follows:

<u>Coefficients:</u>		
	<u>Estimate</u>	<u>Std. Error</u>
(Intercept)	17.227	7.333
Y	0.793	0.298
T	-0.769	0.277
Conf	1.874	0.802

R-squared: 0.8344, Adjusted R-squared: 0.8290		
F-statistic: 118.72 on 3 and 46 DF		
RSS = 1762.88		

(6) a. Test the overall significance of the model at the 5% significance level.

Economists often believe that a reduction on taxes or an increase in income will affect consumption in the same way. In other words, economists often do not include both taxes and income in their models and instead use disposable income (which is income – taxes). This basically forces the effect of a \$1,000 increase in income on consumption to be the same as a \$1,000 decrease in taxes (or that $\beta_1 = -\beta_2$).

So, you change your model to:

$$Cons_i = \beta_4 + \beta_5 Yd_t + \beta_6 Conf_t + \varepsilon_t$$

where:

Yd_t = real disposable income in time t (in millions of dollars)

All other variables are as defined above.

This is a restricted model. The results from this second model are:

<u>Coefficients:</u>		
	<u>Estimate</u>	<u>Std. Error</u>
(Intercept)	17.529	7.927
Yd	0.789	0.290
Conf	1.799	0.801

R-squared: 0.8219, Adjusted R-squared: 0.8127		
F-statistic: 115.68 on 2 and 47 DF		
RSS = 1899.78		

(8) b. Perform a test to see if it is best to use disposable income rather than income and taxes separately in your model. In other words, test the hypothesis that $\beta_1 = -\beta_2$. (Test at the 5% level of significance.)

4. You are estimating a model of the financial aid awards at a local university. Given 50 observations, you estimate the following:

$$FinAid_i = \beta_0 + \beta_1 Parent_i + \beta_2 HSRank_i + \varepsilon_t$$

where:

$FinAid_i$ = financial aid awarded to the i th applicant (in dollars)

$Parent_i$ = amount of parental contribution expected for i th applicant (in dollars)

$HSRank_i$ = high school rank of the i th applicant

The estimated results are:

<u>Coefficients:</u>		
	<u>Estimate</u>	<u>Std. Error</u>
(Intercept)	8927	3298.8
Parent	-0.358	0.032
HS Rank	87.401	20.682

R-squared: 0.7416, Adjusted R-squared: 0.7299		
F-statistic: 36.72 on 2 and 47 DF		

(4) a. Do the coefficients for $FinAid_i$ and $Parent_i$ have the signs you would expect? Explain.

Suppose it is alleged that the financial aid awards are unfairly biased toward males, so you decide to test this claim by adding a gender variable to your model. Your new model is:

$$FinAid_i = \beta_0 + \beta_1 Parent_i + \beta_2 HSRank_i + \beta_3 Male_i + \varepsilon_t$$

where:

$Male_i$ = 1 if applicant i is a male; = 0 otherwise

All other variables are as listed above.

The estimated results are:

<u>Coefficients:</u>		
	<u>Estimate</u>	<u>Std. Error</u>
(Intercept)	9813	3411.6
Parent	-0.337	0.031
HS Rank	83.311	20.098
Male	-1570.3	784.12

R-squared: 0.7605, Adjusted R-squared: 0.7512
F-statistic: 38.56 on 3 and 46 DF

(4) b. Interpret the coefficient on $Male_i$.

(6) c. Test if the accusations are correct – that males receive higher financial aid awards. (Test at the 5% level.)

(8) d. Which of these two models is better? Carefully use the four specification criteria to make your decision and also to explain your answer.

5. For the following models, write out the interpretation of the indicated variable:

(4) a. $Y_i = 1,191 + 2.43X_1 - 3.76X_2$

Interpret the coefficient on X_2 .

(4) b. $\ln Y_i = 955 + 4.77 \ln X_1 + 3.55X_2$

Interpret the coefficient on $\ln X_1$.

(4) c. $\ln Y_i = 2.6 + 0.122X_1 + 0.092X_2$

Interpret the coefficient on X_2 .

(4) d. $Y_i = -46 + 428.3 \ln X_1 - 19.44X_2$

Interpret the coefficient on $\ln X_1$.