**FINM3123 Introduction to Econometrics**

**Test**

**Name: ­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Student ID: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Date: 15 November 2024**

**Time allowed: 120 minutes Full mark: 80**

1. **Multiple choice questions (30 points):**

*Identify the letter of the choice that* ***best*** *completes the statement or answers the question. 2 points for each question.*

1. Which of the following is an example of time series data?

a. Data on the unemployment rates in different parts of a country during a year.

b. Data on the consumption of wheat by 200 households during a year.

c. Data on the gross domestic product of a country over a period of 10 years.

d. Data on the number of vacancies in various departments of an organization on a particular month.

1. The Gauss-Markov theorem will not hold if \_\_\_\_\_.

a. the error term has the same variance given any values of the explanatory variables

b. the error term has an expected value of zero given any values of the independent variables

c. the regression model relies on the method of random sampling for collection of data

d. the independent variables have exact linear relationships among them

1. In a multiple regression model, which of the following is correct regarding the value of the adjusted *R*2?
   1. It can be negative.
   2. It has to be positive.
   3. It has to be larger than the *R*2.
   4. It can be larger than 1.

**Answer questions 4-8 based on the information given in Exhibit 1.**

**Exhibit 1**

The following estimated regression model was developed relating yearly income (Y in $1,000s) of 30 individuals with their age (X1) and their gender (X2) (0 if male and 1 if female).

 = 30 + 0.7X1 + 3X2

Also provided are SST = 1,200 and SSR = 384.

1. Refer to Exhibit 1. From the above function, it can be said that the expected yearly income of

a. males is $3 more than females

b. females is $3 more than males

c. males is $3,000 more than females

d. females is $3,000 more than males

1. Refer to Exhibit 1. The *R*2 is

a. 0.32

b. 0.42

c. 0.68

d. 0.50

1. Refer to Exhibit 1. If we want to test for the significance of the model, the critical value of *F* at 95% confidence is

a. 3.33

b. 3.35

c. 3.34

d. 2.96

1. Refer to Exhibit 1. The test statistic for testing the significance of the model is

a. 0.73

b. 1.47

c. 28.69

d. 5.22

1. Refer to Exhibit 1. The model

a. is significant

b. is not significant

c. would be significant is the sample size was larger than 30

d. None of these alternatives is correct.

1. If j, an unbiased estimator of j, is also a consistent estimator of j,then when the sample size tends to infinity:

a. the distribution of j collapses to a single value of zero.

b. the distribution of j diverges away from a single value of zero.

c. the distribution of j collapses to the single point j.

d. the distribution of j diverges away from j.

1. Which of the following models is used quite often to capture decreasing or increasing marginal effects of a variable?

a. Models with logarithmic functions

b. Models with quadratic functions

c. Models with variables in level

d. Models with interaction terms

1. An interaction term X1X2 in a multiple regression model Y = β0 + β1X1 + β2X2 + β3X1X2 + u is used when
   1. the coefficient of determination is small.
   2. there is a curvilinear relationship between the dependent and independent variables.
   3. neither one of 2 independent variables contribute significantly to the regression model.
   4. the relationship between *X*1 and *Y* changes for differing values of *X*2.

**Answer questions 12-15 based on the information given in Exhibit 2.**

**Exhibit 2**

One of the most common questions of prospective house buyers pertains to the average cost of heating in dollars (Y). To provide its customers with information on that matter, a large real estate firm used the following 4 variables to predict heating costs: the daily minimum outside temperature in degrees of Fahrenheit (), the amount of insulation in inches (), the number of windows in the house (), and the age of the furnace in years (). Given below are the outputs of two regression models.

**Model 1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Regression Statistics* | |  |  |  | |
| R Square | 0.8080 |  |  |  | |
| Adjusted R Square | 0.7568 |  |  |  | |
| Observations | 20 |  |  |  | |
|  |  |  |  |  | |
|  | *Degrees of Freedom* | *Sum of Squares* |  | |
| Regression | 4 | 169503.4241 |  | |
| Residual | 15 | 40262.3259 |  | |
| Total | 19 | 209765.75 |  | |
|  |  |  |  |  | |
|  | *Coefficients* | *Standard Error* | *t Stat* | *p-value* | |
| Intercept | 421.4277 | 77.8614 | 5.4125 | 7.2E-05 | |
| X1 (Temperature) | -4.5098 | 0.8129 | -5.5476 | 5.58E-05 | |
| X2 (Insulation) | -14.9029 | 5.0508 | -2.9505 | 0.0099 | |
| X3 (Windows) | 0.2151 | 4.8675 | 0.0442 | 0.9653 | |
| X4 (Furnace Age) | 6.3780 | 4.1026 | 1.5546 | 0.1408 | |

**Model 2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Regression Statistics* | |  |  |  |
| R Square | 0.7768 |  |  |  |
| Adjusted R Square | 0.7506 |  |  |  |
| Observations | 20 |  |  |  |
|  |  |  |  |  |
|  | *Degrees of Freedom* | *Sum of Squares* |
| Regression | 2 | 162958.2277 |
| Residual | 17 | 46807.5222 |
| Total | 19 | 209765.75 |
|  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *p-value* |
| Intercept | 489.3227 | 43.9826 | 11.1253 | 3.17E-09 |
| X1 (Temperature) | -5.1103 | 0.6951 | -7.3515 | 1.13E-06 |
| X2 (Insulation) | -14.7195 | 4.8864 | -3.0123 | 0.0078 |

1. Referring to Exhibit 2, what can we say about Model 1?
2. The model explains 77.7% of the sample variability of average heating costs; after correcting for the degrees of freedom, the model explains 75.1% of the sample variability of average heating costs.
3. The model explains 75.1% of the sample variability of average heating costs; after correcting for the degrees of freedom, the model explains 77.7% of the sample variability of average heating costs.
4. The model explains 80.8% of the sample variability of average heating costs; after correcting for the degrees of freedom, the model explains 75.7% of the sample variability of average heating costs.
5. The model explains 75.7% of the sample variability of average heating costs; after correcting for the degrees of freedom, the model explains 80.8% of the sample variability of average heating costs.
6. Referring to Exhibit 2 and let α = 5%, what is the decision and conclusion for the test

 using Model 1, where β1, …, and β4 are the coefficients of X1, …, and X4 respectively?

* 1. Do not reject H0 and conclude that the 4 independent variables have significant individual linear effects on average heating costs.
  2. Reject H0 and conclude that the 4 independent variables, taken as a group, have significant linear effects on average heating costs.
  3. Do not reject H0 and conclude that the 4 independent variables, taken as a group, do not have significant linear effects on average heating costs.
  4. Reject H0 and conclude that the 4 independent variables, taken as a group, do not have significant linear effects on average heating costs.

1. Referring to Exhibit 2, what is the value of the *F* test statistic for ?
2. 0.820
3. 1.219
4. 1.382
5. 15.787
6. Referring to Exhibit 2 and let α = 5%, what is the decision and conclusion for the test

?

* 1. Do not reject H0 and conclude that the 2 independent variables have significant individual linear effects on average heating costs.
  2. Reject H0 and conclude that the 2 independent variables, taken as a group, have significant linear effects on average heating costs.
  3. Do not reject H0 and conclude that the 2 independent variables, taken as a group, do not have significant linear effects on average heating costs.
  4. Reject H0 and conclude that the 2 independent variables, taken as a group, do not have significant linear effects on average heating costs.

**II. Problems (50 points)**

1. (12 points)

A research has international cross-sectional data on aggregate wages, *W*, aggregate profits, *P*, and aggregate income, *Y*, for a sample of *n* countries. By definition,

The regression

are fitted using OLS regression analysis. Show that the regression coefficients will automatically satisfy the following equations:

1. (14 points) A commercial real estate company wants to predict market rental rates in a particular metropolitan area, as a service to clients. It aims to do this by fitting a linear model. The response variable of interest is rental rate (rent), and the other variables are predictors, including age (age), operating expenses and taxes (expense), vacancy rates (vacancy), and total square footage (space). The dataset contains suburban commercial properties. The output is shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. (6 points) For this model, carry out the test of against (2-sided / 2-tailed test) at the 5% significance level.

A simpler model without expense and vacancy is also investigated. The output is shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. (8 points) Test at the 5% significance level.
2. (12 points)
3. (3 points) Recall the definition of consistency of an estimator.
4. (3 points) Consider the simple regression model . The fitted regression line is given by  . What are the assumptions needed about this model to ensure that is a consistent estimator of ?
5. (6 points) Under the same set of assumptions, use the consistency of to prove mathematically that is a consistent estimator of . [*Hint*: use the law of large numbers. ]
6. (12 points) Consider the regression model with a sample size *n*. Show that the adjust R-squared increases, if and only if, the t-statistic of a newly added regressor is greater than one in absolute value. [Hint: If a random variable , where *tm* denotes a *t* distribution with *m* degrees of freedom, then , where is an F distribution with 1 numerator degree of freedom and *m* denominator degrees of freedom.]



