

Name \_\_\_\_\_

Student No. \_\_\_\_\_

### STA 442/1008 Quiz 9

1. (5 Points) In a study of adaptation to urban life in Peru, a sample of women living in the city were classified according to a variable we shall call "Origin." They were either

- 1 = Aboriginal ancestry, born at high altitude
- 2 = Aboriginal ancestry, born at low altitude
- 3 = Hispanic ancestry, born at low altitude

Age, income and obesity (as indicated by percentage of body fat) were also recorded for each participant in the study. Consider the following full (unrestricted) regression model.

$$E[Y|X] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 D_1 + \beta_4 D_2, \text{ where}$$

$Y$  = Obesity

$X_1$  = Age

$X_2$  = Income

$D_1$  = 1 if Origin = 1, zero otherwise

$D_2$  = 1 if Origin = 2, zero otherwise

and each conditional distribution of  $Y$  is normal with the same variance  $\sigma^2$ .

Give the **null hypothesis** you would test in order to answer the following questions. Note that a null hypothesis specifies restrictions on the regression parameters, yielding a reduced (restricted) model. In each case, the answer is equivalent to one or more linear combinations of  $\beta$  values set equal to zero.

- a. Controlling for Age and Income, is Obesity related to Origin?
- b. Controlling for Origin, is either Age or Income (or both) related to Obesity?
- c. Controlling for Age and Income, is there a difference in average Obesity level between Aboriginals born at low altitude and Hispanics born at low altitude?
- d. Controlling for Age and Income, is there a difference in average Obesity level between Aboriginals born at low altitude and Aboriginals born at high altitude?
- e. Controlling for Age and Income, is there a difference between the average Obesity of Hispanics born at low altitude and the mean of Aboriginals born at high altitude and Aboriginals born at low altitude?

**Continued on Reverse**

2. (5 Points) The questions below refer to your analysis of the TV data. For each question, give the numerical value of the test statistic and state whether the relationship is statistically significant at the 0.05 level.

- a. Controlling for total number of people in the household and assessed value of home, is there a difference between Rural and Urban locations in average number of TV sets?

Numerical value of $F$ or $t$ statistic	Statistically significant? (Yes or No)

- b. Controlling for total number of people in the household and assessed value of home, is there a difference between Small Town and Urban locations in average number of TV sets?

Numerical value of $F$ or $t$ statistic	Statistically significant? (Yes or No)

- c. Controlling for total number of people in the household and assessed value of home, is there a difference between Rural and Small Town locations in average number of TV sets?

Numerical value of $F$ or $t$ statistic	Statistically significant? (Yes or No)

- d. Controlling for location and assessed value of home, is total number of people in the household related to average number of TV sets?

Numerical value of $F$ or $t$ statistic	Statistically significant? (Yes or No)

- e. State the finding for (d) in simple, non-statistical language.

**Total Marks = 10 points**