

Name Jerry

Student Number \_\_\_\_\_

## STA 442/2101 f2014 Quiz 7

1. (6 points) In a study comparing the effectiveness of different weight loss diets, volunteers were randomly assigned to one of two diets (A or B) or put on a waiting list and advised to lose weight on their own. Participants were weighed before and after 6 months of participation in the program (or 6 months of being on the waiting list). The response variable is weight loss. The explanatory variables are age (a covariate) and treatment group.

- (a) Write the regression equation. Your model should have *no intercept*, and *parallel regression lines*. Please use  $x$  for age. You don't have to say how your dummy variables are defined. You'll do that in the next part.

$$Y_i = \beta_1 d_{i1} + \beta_2 d_{i2} + \beta_3 d_{i3} + \beta_4 x_i + \varepsilon_i$$

okay to give  $E(Y|x, d)$

- (b) Make a table with three rows, showing how you would set up indicator dummy variables for treatment group. Give  $E(Y|x)$  in the last column.

	$d_1$	$d_2$	$d_3$	$E(Y x)$
A	1	0	0	$\beta_1 + \beta_4 x$
B	0	1	0	$\beta_2 + \beta_4 x$
Control	0	0	1	$\beta_3 + \beta_4 x$

- (c) In terms of  $\beta$  values, what null hypothesis would you test to find out whether, allowing for age, the three diets (including Wait List) differ in their effectiveness?

$$H_0: \beta_1 = \beta_2 = \beta_3$$

- (d) In terms of  $\beta$  values, what null hypothesis would you test to find out whether, allowing for age, diets A and B differ in their effectiveness?

$$H_0: \beta_1 = \beta_2$$

- (e) In terms of  $\beta$  values, what null hypothesis would you test to find out whether the Wait list "diet" is of any value at all in helping 25-year-old participants to lose weight? Remember,  $Y$  is weight loss, which could be zero or even negative.

$$H_0: \beta_3 + 25\beta_4 = 0$$

- (f) Is it safe to assume that age is independent of treatment group? Answer Yes or No and briefly explain.

Yes, because of random assignment

2. These questions are based on your analysis of the Census Tract data. **If you do not have a number on your printout, do NOT write it on your quiz paper.** It will be assumed that you copied it from somebody, and you will be charged with an academic offense. Also, you can't get credit for conclusions without the numbers.

(a) (2 Points) You tested for a difference in expected crime rate between the Northeast and West regions, after controlling for all the other variables in the model.

- i. What is the value of the test statistic? (Not the  $p$ -value, the test statistic.) The answer is a number from your printout. Write the number below, and also circle it on your printout. On the printout, write "Question 2a" beside the number.

$$t = 5.107$$

- ii. Guided by the  $\alpha = 0.05$  significance level as usual, what do you conclude? The answer is something about expected crime rate. You must state a *directional* conclusion if one is possible.

Controlling for other variables in the model, expected crime rate is greater in the West than the Northeast.

(b) (2 Points)

- i. Controlling for Region and Percent High School Graduates, you tested all the other variables at once. What is the value of the test statistic? (Not the  $p$ -value, the test statistic.) The answer is a number from your printout. Write the number below, and also circle it on your printout. On the printout, write "Question 2b" beside the number.

$$F = 2.5455$$

- ii. Guided by the  $\alpha = 0.05$  significance level as usual, what do you conclude? Any kind of language (technical or not) is okay as long as it's possible to understand what you mean.

Controlling for Region and Percent High School graduates, at least one of the other variables is related to crime rate.

Please attach your log file and your list file. Make sure your name is on both files.