*YOUR TA’S NAME*:

*Lecture Worksheet*

*Tuesday 12/8/2020*

**MAIN POINTS OF LECTURE**

1. TOPIC #1: In a model with a discrete independent variable X that has j categories, X should be represented by a series of j-1 “dummy variables” that indicate whether individuals belong to categories of X.

This is directly analogous to ANOVA.

1. TOPIC #2: Interaction terms … a strategy for allowing the effect of X1 on Y to vary across levels of X2 and simultaneously allowing the effect of X2 on Y to vary across levels of X1 … can be modeling by adding a new variable that equals X1×X2.

When we do this, X1 moderates the effect of X2 on Y and (simultaneously) X2 moderates the effect of X1 on Y

**QUESTIONS**

**From the recorded lecture**

1. The prediction equation below is from the regression of continuous variable Y (income) on discrete variable X (highest degree attained). Note that X1=1 if people did not complete high school (and 0 otherwise).

$$\hat{Y}=20,000+15,000X\_{2}\left(High School Diploma Only\right)$$

$$+25,000X\_{3}\left(Bachelors Degree Only\right)$$

$$+45,000X\_{4}(Advanced Degree)$$

Report the mean value of Y for each of the 4 discrete values of X. That is, what is the mean level of income for people with different levels of education?

1. The prediction equation below is from a regression of continuous variable Y (“happiness” where 100=maximum happy and 0=maximum unhappy) on continuous variable X1 (“age in years”); discrete variable X2 (0=Packers fan, 1=Vikings fan); and discrete interaction term X3 which equals X1 times X2.

$$\hat{Y}\_{i}=10+1.0X\_{1}(Age)-10.0X\_{2}(Fan)-0.5X\_{3}$$

* 1. How do you interpret the coefficient (aka, slope) for X1?
	2. What is the effect of age for Vikings fans?
	3. What is the effect of age for Packers fans?