

LAST NAME (Print): \_\_\_\_\_ FIRST: \_\_\_\_\_

Statistics 111      **Quiz 20**

1. Multiple regression on 18 students predicts final grade (FG) from the hours spent studying the night before the final (NB), the hours spent studying the week before the final (WB), and the number of courses being taken (NC).

$$FG = 1.2 - 0.3 * NB + 0.2 * WB + 0.3 * NC$$

The residuals have sd 0.6. The se's for the NB, WB, and NC coefficients are 0.1, 0.05, 0.2, respectively.

\_\_\_\_\_ You study two hours the night before, ten hours the week before, and take three classes. What grade do you expect?

\_\_\_\_\_ You test whether NC is useful. What is the value of your test statistic?

\_\_\_\_\_ What distribution do you use? (Include df if appropriate.)

\_\_\_\_\_ Should the variable be used in the model? Use  $\alpha = .05$ .

2. You predict height from shoe size. With a sample of 9, the regression equation is  $Y = 2 + X/3$ . The sd of the residuals is 0.5, and the mean and variance of the shoe sizes are 8 and 7, respectively. The coefficient of determination is 0.7.

\_\_\_\_\_ What is your prediction for the height of someone with size 10 shoes?

Set a two-sided 95% confidence interval on the average height of people with size 10 shoes.

L = \_\_\_\_\_      U = \_\_\_\_\_

3. List all, and only, true statements. \_\_\_\_\_

- A. If a scatterplot looks like a cigar, then it is heteroscedastic.
- B. Multicollinearity arises when two or more of the explanatory variables are highly correlated.
- C. The response variable is sometimes called the independent variable.
- D. As points cluster more tightly around a regression line, the correlation increases.
- E. If one include irrelevant explanatory variables in the model, it reduces predictive accuracy.
- F. A residual is the observed value minus the predicted value.