

## Homework #5

due Friday, March 29 in class

**Instructions:** Complete all problems and turn in a set of answers either to me or under my door (office 335L) by the assigned due date. Do ask me questions via email or in my office hours. Do work together. Do not copy answers from another student or turn in answers that are substantively identical. To clarify, if you work with another student, I would expect that your numerical answers would be quite close, but that your verbal explanations would be similar, but not identical, reflecting that you wrote up your answers independently. Show your work, and write out explanations for your answers. If you use Excel or a similar tool, write “According to Excel, ...” in your answer.

**Problem 1** Fire damage in the United States amounts to billions of dollars each year, much of it insured. The time taken for firefighters to arrive at the fire is critical. This raises the question, should insurance companies lower premiums if the home to be insured is close to a fire station? To help make a decision, a study was undertaken wherein a number of fires were investigated. The distance to the nearest fire station (in miles) and the percentage of fire damage were recorded (dataset Xr16-11). Determine the least squares regression line and interpret the coefficients.

**Problem 2** An economist wanted to investigate the relationship between office rents (the dependent variable) and vacancy rates. Accordingly, he took a random sample of monthly office rents and the percentage of vacant office space in 30 different cities (dataset Xr16-16).

- a. Determine the regression line.
- b. Interpret the coefficients

**Problem 3** Answer the following questions related to the 2008 General Social Survey (available in the zip file I emailed to you under ANESS-GSS → Excel → GSS → GSS2008.xls).

- a. Conduct an analysis of the relationship between income (INCOME) and age (AGE). Estimate with 95% confidence the average increase in income for each additional year of age. (hint: you may need to drop observations from the dataset for which you are missing information on either income or age)
- b. Is there sufficient evidence to conclude that more educated people (EDUC) watch less television (TVHOURS)?
- c. What does the dataset say about the relationship between education and income?

**Problem 4** A computer dating service typically asks for various pieces of information such as height, weight, and income. One such service requests the length of index fingers. The only plausible reason for this request is to act as a proxy on height. Women in particular have often complained that men lie about their heights. If there is a strong relationship between heights and index fingers, the information can be used to “correct the false claims about heights. To test the relationship between the two variables, researchers gathered the heights and lengths of index fingers (in centimeters) of 121 students (dataset Xr16-107).

- a. Using a computer, draw a scatter plot depicting the relationship between the two variables.
- b. Is there sufficient evidence to infer that height and length of index fingers are linearly related?

c. Predict with 95% confidence the marginal increase in height associated with a 1cm increase in index finger length.

**Problem 5** A researcher is interested in the effect of the number of police officers working in a city on that city's violent crime rate. She obtains data on the 500 largest cities in the US, with the 2012 violent crime rate and the number of police officers per capita employed in 2012 for each city. She then uses a computer to regress the violent crime rate on police per capita, and finds the following results:

$$R^2 = .64$$

	Coefficients	Standard error	t stat	P-value
Intercept	604.32	158.65	3.81	$1.39 * 10^{-4}$
Police per capita	22.13	10.43	2.12	.034

a. Interpret the regression results. What do the results predict is the effect of hiring one more police officer? Is this effect significant?

b. A colleague points out that the positive coefficient estimated for the police variable suggests that police cause crime. Can you suggest an alternative explanation?

**Problem 6** An economist for the federal government is attempting to produce a better measure of poverty than is currently in use. To help acquire information, she recorded the annual household income (in thousands of dollars) and the amount of money spent on food during one week for a random sample of households (dataset Xr16-15).

a. Determine the regression line and interpret the coefficients.

b. Determine the coefficient of determination and describe what it tells you.

c. Conduct a test to determine whether there is evidence of a linear relationship between household income and food budget.