Possible Solutions for Homework 11

Econ 29000

Kevin R Foster, CCNY

This exercise will use the version of the CPS dataset online, CPS_hw11, which is a version of the 2010 March CPS data from the US Bureau of Labor Statistics. (It is a zipped file so first download it and then unzip before you open it with SPSS.) It collects a variety of information, which you are welcome to look through as you consider final project topics.

Use "Data\Select Cases" to choose only people with Total Wage and Salary greater than zero.

1. Please list the names of the people in your study group.
2. Use SPSS to run a simple linear regression (choose Analyze\Regression\Linear) with the dependent variable as Total Wages and Salary and the independent variable Age.
    a. What is the estimated value for β_0? Is it statistically significantly different from zero?
    b. What is the estimated value for β_1? Is it statistically significantly different from zero?

The SPSS output is:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	12698.652	489.894		25.921	.000
	Demographics, Age	716.714	11.411	.196	62.809	.000

a. Dependent Variable: Total wage and salary earnings amount - Person

The β0 (beta-zero or beta naught) coefficient estimate is $12,698.65 – this implies that people are predicted to start from just under $13,000 per year when born and then add wages as they age. It is statistically significantly different from zero: the standard error is just 490, so the coefficient estimate is almost 26 times larger than the standard error. The p-value is less than 0.000 (certainly less than 0.05).

The β1 (beta-one) coefficient estimate is $717 – implying that people get an average of $717 more in wages & salary for each year that they get older. This is also statistically significant; its standard error is just 11; the coefficient estimate is almost 63 times larger. The p-value is nearly zero, certainly less than 0.05.

c. What is the predicted value for someone who is 25 years old? 45?

The predicted wage is $12,699 + 717*Age so someone who is 25 is predicted to earn $30,607; someone 45 years old would be predicted to earn $44,951.

d. If you had not selected only people with non-zero wages, how would that change the estimates? (You can undo the "select cases" and see for yourself.)
e. How different would the estimates be, if you had selected people with non-zero wages and also who were "prime-age" so 25-55 years old?
f. Again use "select cases" to do a regression on just prime-age women. Then do a separate regression for prime-age men. What are the differences?

These are shown in the table below.

	wage > 0	all	prime age	prime age men	prime age women
β_0	12699	6675	16404	13015	19631
β_1	717	376	752	1068	428

Clearly adding in all of the zero wages drags down both coefficient estimates. People at "prime age" have a similar slope but a higher level of wages (suggestive of curvature); men start lower than women but accrue increases at a faster rate.

3. Next estimate a linear regression with the same dependent variable, but now add the "female" dummy variable and education variables (you can choose which ones). Interpret the regression coefficients: which are statistically significant? (You can choose whether to do prime-age or all ages, but you'll need both men and women.)

These estimates are:

	wage>0	prime age
constant	7594	2809
Age	495	723
Female	-18598	-21275
just HS	11018	10350
some college w/o degree	15630	17961
Assoc in vocational	20702	20050
Assoc in academic	23491	23088
4-yr degree	38798	39500
Adv degree	65299	65435

Women are predicted to do even worse if I just focus on prime-age; the educational premiums are quite close for both groups however.