|  |
| --- |
| By Cliff Rodriguez |
| Problem Set 1 |
| Due: January 31, 2018 |

|  |
| --- |
| Cliff Rodriguez |

Contents

[Question 1 2](#_Toc504427787)

[Question 2 2](#_Toc504427788)

[Question 3 2](#_Toc504427789)

[Question 4 3](#_Toc504427790)

**No table of figures entries found.**

# Question 1

1. is a constant that represents a value for the impact of non-internalized variables that are relevant to the system.

From the class notes: error term or random disturbance

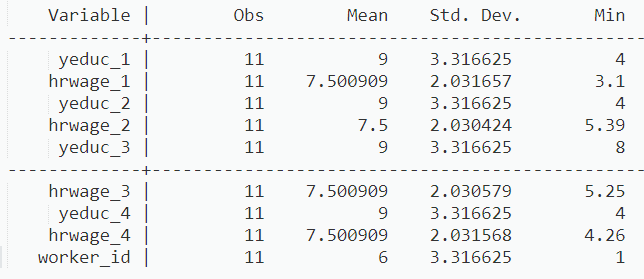
another way to think ofuiit is all the other stuff that impacts Y but is notincluded in the modelmany timesuiis thought of as theunobservedfactors that impact Ywe never know or observeui

1. Three examples of things that could be included in for this model are
   1. Height of father
   2. Heightof Mother
   3. Were prenatal vitamins used by the mother
2. Give and intuitive argument for why:
   1. might be positive in this model if
   2. might be negative in this model if
3. Simple regression a constant

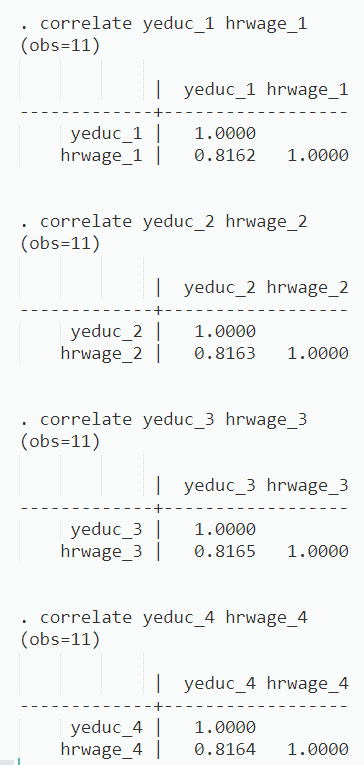
# Question 2

* 1. The mean and standard deviation for each variable in dataset ps1q2.dta are presented below.

*Calculated using the command sum*



* 1. The covariance and correlation between years of education and hourly wages in dataset ps1q2.dta are presented below.



1. For each firm the OLS regress for hourly wagers on years of education is below.
2. Based on calculations made using the OLS method, for each firm the added value from one year of education is XX
3. The relationship between hourly wages and years of education for each firm is graphed below.
4. Reviewing the graphs the prediction is (not?) equally good for each firm. This is because…
5. Reviewing the graphs the relationship between hourly wages and years of education is/is not the same? This is because…
6. Based on points *v* and *vi* above it is suggested that ……. About the choice of model in estimation

# Question 3

1. See appendix I for work

*=*

*=*

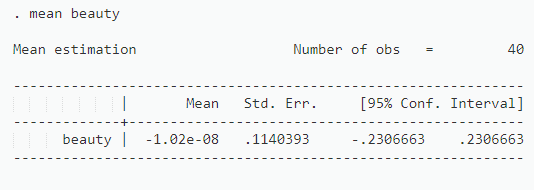
1. The term in this model is/is not useful because…
2. The GPA score is predicted to be …. if the ACT score increases by 5 points?
3. The fitted values and residuals for each observation are presented in appendix II.

Calculation showing residuals approximately sum to zero:

1. Blank of the variation in GPA for the eight students is explained by the ACT. This is because…

# Question 4

1. *=*



1. *=*
2. The covariance between course evaluation and beauty is …

The units of measure for the covariance between course evaluation and beauty is … and this does/does not have a real world interpretation.

1. The correlation between course evaluations and beauty is shown below using STATA and the

1. The data with beauty plotted on the x-axis is below.
2. Using to calculate the regression slope coefficient the value is …
3. For the data in ps1q4.dta the value of is

This does or does not relate to *course\_eval* because

1. Using the regression

*The OLS estimates are:*

*compare in step vi with found above:*

1. *The*  measures

The standard error (RMSE) of a regression measures…

I prefer R2 or RMSE? WHY?

1. *in step vii is*