EC 320: Introduction to Econometrics
Lecture Notes

David M. Hall
University of Oregon

Fall 2025

1 Session Overview

- Date: October 7, 2025
- Topic: Regression Output: Intrerpretation, R^2 and Exogeneity
  - Chapters (Dougherty): 1.4-1.6
- Goals/Objectives:
  1. Interpret the Output of a Simple Linear Regression
  2. Derive the Exogeneity Assumption(s)
  3. Understand the importance of and derive R^2 metric.

1.1 Before Class:

2 During Class:

2.1 Welcome

- Explain goals and plan for the day.

2.1.1 Warm Up:

Example 2.1. Problem Set Problem

2.2 Content Section #1:

Time: 15 Minute Review

In the last lecture, we introduced the simple linear regression model. We supposed that we are regressing a variable, Y (our dependent variable) on a single regressor X (independent/explanatory/regressor). We are going over a very basic way of determining the relationship between these two: that of a simple **linear** regression, which looks as follows

$$Y_i = \beta_1 + \beta_2 X_i + u_i \tag{1}$$

The β's are coefficients on the model, Y_i and X_i are observations (data), and u_i is a disturbance term.

What we will be using in this course is the **least squares algorithm**. Recall that the steps are as follows:

1. Isolate the residual.
2. Square the residual.
3. Sum the squared residuals (RSS).
4. Take partial derivatives of b_0 and b_1.
5. Solve the system of equations for b_0 and b_1.

We were able to do this for the general least squares algorithm, where you have some n observations. Recall that we were fitting:

$$\hat{Y_i} = b_0 + b_1 X_i \tag{2}$$

We can then define the residual:

$$\hat{u_i} = Y_i - \hat{Y_i} = Y_i - b_0 - b_1 X_i \tag{3}$$

which holds for **all** observations i of our total observations n. Squaring this, we get:

$$\hat{u_i^2} = Y_i^2 + b_0^2 + b_1^2 X_i^2 - 2b_0 Y_i - 2b_1 X_i Y_i + 2b_0 b_1 X_i \tag{4}$$

which, again, holds for **all** i. We then sum this over all i in n using our summation operator $\sum_{i=1}^{n}$ which is a *linear* operator.

$$RSS = \sum_{i=1}^{n} \hat{u_i^2} = \sum_i Y_i^2 + n \cdot b_0^2 + b_1^2 \sum_i X_i^2 - 2b_0 \sum_i Y_i - 2b_1 \sum_i X_i Y_i + 2b_0 b_1 \sum_i X_i \tag{5}$$

taking our partial derivatives:

$$\frac{\partial \text{RSS}}{\partial b_1} = 2nb_1 + 2b_2 \sum X_i - 2 \sum Y_i = 0 \tag{1}$$

$$\frac{\partial \text{RSS}}{\partial b_2} = 2b_2 \sum X_i^2 + 2b_1 \sum X_i - 2 \sum X_i Y_i = 0 \tag{2}$$

We then showed that, since this is a system of **two** equations with **two** unknowns, we can solve for b_0 and b_1. Working with equation (1) and noting that the mean is defined as $\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$, we get that:

$$b_0 = \bar{Y} - b_1\bar{X}$$

which we can substitute into equation (2) to get:

$$\boxed{b_2 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}}$$

~ Questions? ~

So we know how to obtain the estimates for the functional relationship between Y and X in a simple linear regression... but **what do they mean?**

To think about this, I'd like to provide you with two quick examples in R. Both use made up data. The first is a regression of exam outcomes on hours of study, the second is a regression of exam outcomes on hours of sleep.

The following code can be used to generate the dataset in R, with the module in Canvas containing the entirety of the code:

```
```{r exam scores df}
    N <- 70 # number of students
    beta_study <- .1 # effect of study hour
    beta_sleep <- 0.05 # effect of sleep hour
    set.seed(320) # for replication

    # creating a data frame
    exam_scores <- tibble(
      i = 1:N,
      study_hours = rnorm(n = N, mean = 4, sd = 2),
      sleep_hours = rnorm(n = N, mean = 7, sd = 1),
      u = rnorm(n = N, mean = 0, sd = 0.05) # random error
    )

    # ensure no negative hours of study or sleep
    exam_scores <- exam_scores %>%
      mutate(
        study_hours = ifelse(study_hours < 0, 0, study_hours),
        sleep_hours = ifelse(sleep_hours < 0, 0, sleep_hours),
        exam_score = 0.01 + study_hours*beta_study + sleep_hours * beta_sleep + u
      )
```
```

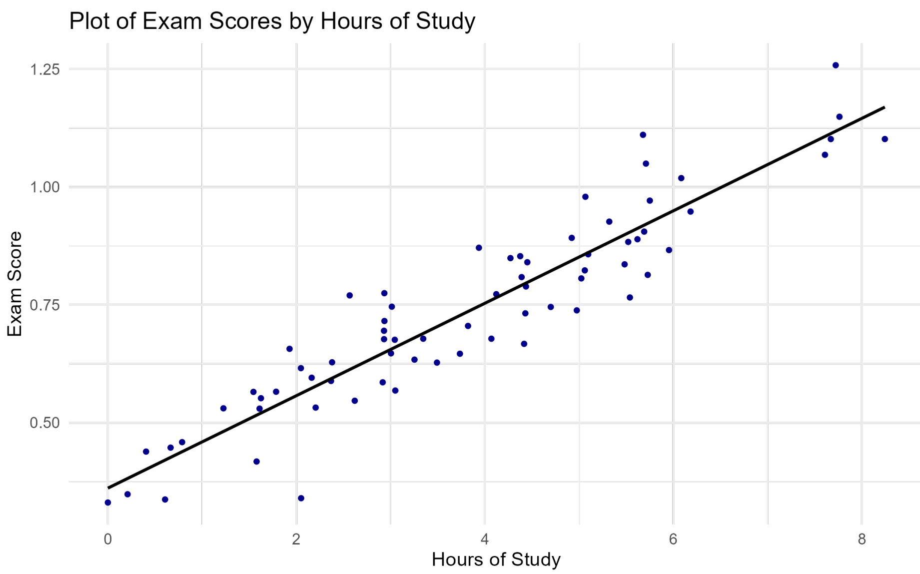


Table 1: Regression Results of Exam Scores on Hours of Study

	Dependent variable:
	Exam Score Study Hours
study_hours	0.098***
	(0.005)
Constant	0.361***
	(0.020)
Observations	70
R^2	0.863
Adjusted R^2	0.861
Residual Std. Error	0.077 (df = 68)
F Statistic	429.766*** (df = 1; 68)
Note:	*p<0.1; **p<0.05; ***p<0.01

Let's take a look at the regression output of our simple linear regression of exam scores along with a chart of the fitted line, one for each of our figures.

The way we interpret effectively any linear regression so think of the coefficient on your Xvariable as a "one unit" change. We measure the units in Xas well. So in Table 2.2, since we are regressing exam score on study hours, we interpret the slope coefficient as the effect on exam score of a **one hour increase** in study time.

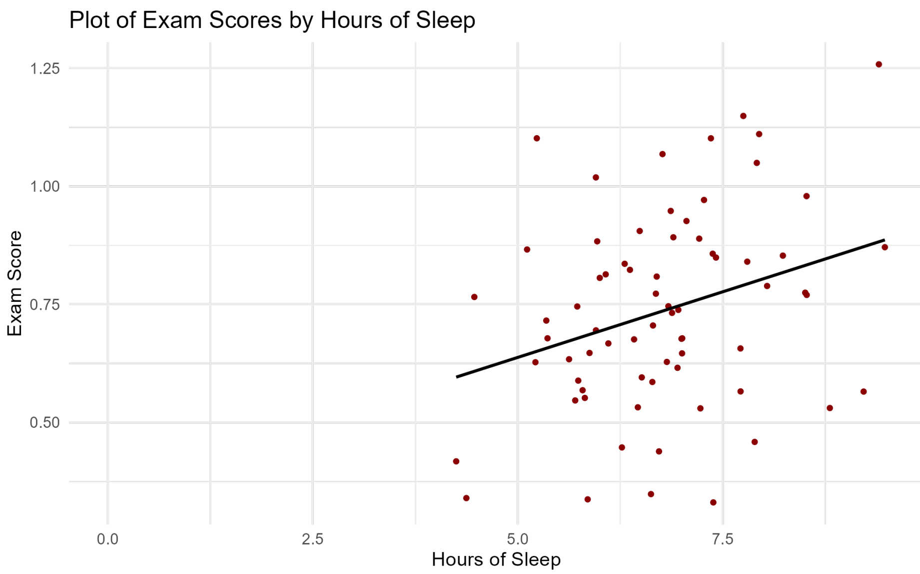


Table 2: Regression Results of Exam Scores on Hours of Sleep

	Dependent variable:
	Exam Score Sleep Hours
sleep_hours	0.056**
	(0.021)
Constant	0.360**
	(0.145)
Observations	70
R^2	0.092
Adjusted R^2	0.079
Residual Std. Error	0.200 (df = 68)
F Statistic	6.922** (df = 1; 68)
Note:	*p<0.1; **p<0.05; ***p<0.01

In Table 2.2 I show the output of the regression of exam scores on hours of sleep. Same thing here: this is telling us the effect of an increase of one hour of sleep on exam score.

One way to write the interpretation verbatim would be as follows:

"For every additional hour of studying for the exam, I/we expect the exam score to increase by 9.8%."

"A one hour increase in the hours of sleep is estimated to result in an increase of 5.6% on the exam."

2.2.1 Interpreting the Intercept

What about the intercept? The way we interpret the intercept is the value of Y if our X were 0. So in this case, what if you did not sleep? What if you did not study at all? The intercept tells us what we would expect your score to be.

In this case, if you did not study at all I would expect your exam score to be 36.1%, and if if you did not sleep at all I would expect you to get a 36%.

Maybe this seems a bit ridiculous — surely if I showed up to every class but just did not study outside of it, the professor will give me enough content to still get at least a D+? And I *never* sleep, I **live** on caffeine![1]

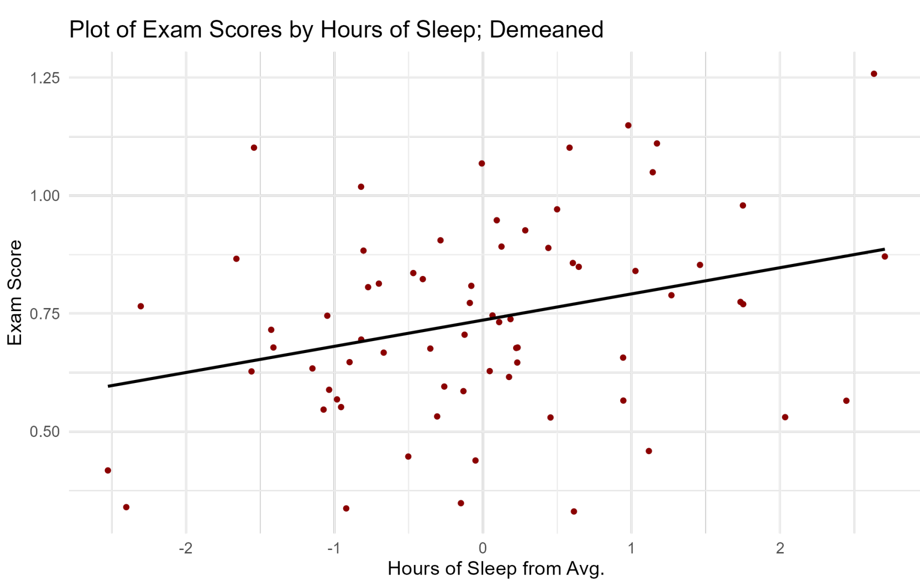
We see in the latter example's plot that most of the data on hours of sleep is *away* from 0. If you, on average, had zero hours of sleep I would not expect you to still get a 36% on any exam. I would expect you to get zero, because you would probably be dead.

What we can do when we expect that the intercept does not have much of an interpretation, because there is no data near where $X = 0$, is to *demean*. In doing so, the slope will not be affected but the *intercept* will be that of the person with the average X.

To demean, instead of regressing Y on X we regress Y on X **demeaned** s.t.

$$X^* = X_i - \bar{X} \tag{6}$$

In doing so, you would get the following output for sleep.



[1] Don't do this.

Table 3: Regression Results of Exam Scores on Hours of Sleep

	Dependent variable:	
	Exam Score	
	Sleep Hours	Sleep Hours Demeaned
	(1)	(2)
sleep_hours	0.056^{**}	
	(0.021)	
sleep_star		0.056^{**}
		(0.021)
Constant	0.360^{**}	0.736^{***}
	(0.145)	(0.024)
Observations	70	70
R^2	0.092	0.092
Adjusted R^2	0.079	0.079
Residual Std. Error (df = 68)	0.200	0.200
F Statistic (df = 1; 68)	6.922^{**}	6.922^{**}
Note:		*p<0.1; **p<0.05; ***p<0.01

We can see now that the actual coefficient on the number of sleep hours is unchanged, but the intercept is now that of the average number of hours of sleep rather than when somebody gets no sleep.

~ Questions? ~

2.3 Activity #1:

Take a good look at the output for the number of study hours. Note that the dependent variable is in *hours*, but that we can measure time differently. Suppose, for example, I asked you to tell me what the effect of an additional *minute* of studying was on your exam score? What do you think it would be?

- What is the effect of an additional *minute* on your exam score?
- What about 20 minutes?

Answer:

Sometimes we have data in different units than we want to interpret. For example, you could have data in degrees Celsius but want to communicate in degrees of Fahrenheit.

While it is intuitive, just note that a change in the unit of measurement, as long as it a linear transformation, changes the *coefficient* by the same transformation.

In our examples above, since one minute is $\frac{1}{60}$ of one hour, the coefficient on an additional minute of study is $\frac{1}{60}$ that of an hour, so an increase of 0.163%.

For 20 minutes, we can think of this as $\frac{1}{20}$ of one hour or 20× one minute. Either way, an additional 20 minutes of studying based on the regression should result in 3.27% increase in your exam score.

2.4 Content Section #2:

Time:

Please note that I am skipping over Chapter 1.5 for now, it may be good to read but we will come back to it actually in next chapter anyhow.

In lots of contexts, the goal of a regression is to be able to explain all of the variation in Y. I will next introduce the concept of total sum of squares, which is one way that we measure the total variation in Y:

Definition 2.1 (Total Sum of Squares)**.** $\text{TSS} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$

If Time Allows

There are two results that, if time, we will go over.

1. The mean value of the residuals is zero.
2. The sample correlation between the observations on $X$and residuals is zero.

Recall: the definition of residual:

$$\hat{u}_i = Y_i - \bar{Y} \tag{7}$$

where $Y_i = \beta_0 + \beta_1 X + u_i$ and $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$.
We want to show that the mean value of the residuals is zero. By the definition of the mean:

$$\frac{1}{n}\sum_i \hat{u}_i = \frac{1}{n}\sum_i Y_i - \frac{1}{n}\sum_i \hat{\beta}_0 + \hat{\beta}_1 \sum_i X_i = \bar{Y} - \frac{1}{n}\cdot n \cdot \hat{\beta}_0 - \hat{\beta}_1 \bar{X}$$

and by the definition of $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ then:

$$\bar{\hat{u}} = \bar{Y} - (\bar{Y} - \hat{\beta}_1 \bar{X}) - \hat{\beta}_1 \bar{X} = 0$$

We can also show that the sample correlation of $X$and the residuals is 0.

$$WTS : \sum_i X_i \hat{u}_i = 0$$

$$\hat{u}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i \implies \sum_i X_i \hat{u}_i = \sum_i X_i Y_i - \hat{\beta}_1 \sum_i X_i - \hat{\beta}_2 \sum_i X_i^2$$

which is zero by the least squares algorithm (i.e. the "normal equation" step when summing the square residuals)

If there is no time to explain, I just want to write two quick theorems:

Theorem 2.1 (Mean of Residuals is 0)**.** $\bar{\hat{u}} = 0 \implies \bar{Y} = \bar{\hat{Y}}$

which means that the average of the residuals is zero.

Theorem 2.2 (Orthogonality of Residuals)**.** $\sum_i X_i \hat{u}_i = 0$

there is no correlation between observed values of $X$and the residual. **Note:** this also means that $\sum_i \hat{Y}_i \hat{u}_i = 0$ as well.

What we will next show is that our total sum of squares can be *decomposed* into two parts: the explained portion and the unexplained portion.

Theorem 2.3 (TSS Decomposition)**.** *The total variation in the outcome variable Y can be decomposed into the explained and unexplained parts:*

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{n}\hat{u}_i^2$$

That is, $TSS = ESS + RSS$.

Proof. We start by noting that after regressing Y on X, each observation can be written as:

$$Y_i = \hat{Y}_i + \hat{u}_i$$

Then the total sum of squares is:

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}\left((\hat{Y}_i + \hat{u}_i) - \bar{Y}\right)^2$$

Since $\bar{Y} = \bar{\hat{Y}}$ and $\bar{\hat{u}} = 0$, we can rewrite:

$$= \sum_{i=1}^{n}\left((\hat{Y}_i - \bar{\hat{Y}}) + (\hat{u}_i - \bar{\hat{u}})\right)^2 = \sum_{i=1}^{n}\left((\hat{Y}_i - \bar{\hat{Y}}) + \hat{u}_i\right)^2$$

Now apply the identity $(a+b)^2 = a^2 + 2ab + b^2$:

$$= \sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2 + 2\sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})\hat{u}_i + \sum_{i=1}^{n}\hat{u}_i^2$$

The middle term vanishes because the residuals are orthogonal to the fitted values:

$$\sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})\hat{u}_i = 0$$

Therefore:

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{n}\hat{u}_i^2$$

$$\Rightarrow \text{TSS} = \text{ESS} + \text{RSS}$$

□

What is this saying?!?:

Mathematically, we are noting that the entirety of the variation in the values of Y can be decomposed into that which can be explained, the ESS — the sum of our fitted values minus the average of the value of observed Y — and the unexplained portion RSS which is just the squared sum of the errors.

We now define R^2:

Definition 2.2 (R-squared)**.** R^2 is the proportion of the total sum of squares that is explained in the regression:

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2}{\sum_i (Y_i - \bar{Y})^2}$$

or equivalently:

$$R^2 = 1 - \frac{RSS}{TSS} = \frac{\sum_i \hat{u_i}^2}{\sum_i (Y_i - \bar{Y})^2}$$

Put in words, this is the amount/proportion of the total variation that is explained. The higher the value, the better fit of the entire model.

Let's take a closer look at our regression output, and see how we did fitting the total variation in the R^2:

Table 4: Regression Results of Exam Scores on Hours of Study

	Dependent variable:
	Exam Score Study Hours
study_hours	0.098***
	(0.005)
Constant	0.361***
	(0.020)
Observations	70
R^2	0.863
Adjusted R^2	0.861
Residual Std. Error	0.077 (df = 68)
F Statistic	429.766*** (df = 1; 68)
Note:	*p<0.1; **p<0.05; ***p<0.01

Remark 2.1. One thing I want to note about R^2: many people overemphasize the role that R^2 plays, particularly when they are trying to test the effect of just one variable! The R^2 tells me all about the variation completely in the value of Y, but if I only care about the coefficient on a single variable X, even if it doesn't explain *all* of the variation in Y, then I might be caught thinking that a low R^2 means that I did not do a very good regression. This is not necessarily true.

For example: suppose I want to test the effect of parent income on test scores. My test scores might be dependent on other things such as sleep or my study time, but if all I care about is parent income and I do not add in study time, then my R^2 will be lower. I did not add in a variable that could explain some variation in Y. That does not necessarily mean that my coefficient on parent income is not believable — it might be!

Table 5: Regression Results of Exam Scores on Hours of Sleep

	Dependent variable:
	Exam Score Sleep Hours
sleep_hours	0.056**
	(0.021)
Constant	0.360**
	(0.145)
Observations	70
R^2	0.092
Adjusted R^2	0.079
Residual Std. Error	0.200 (df = 68)
F Statistic	6.922** (df = 1; 68)
Note:	*p<0.1; **p<0.05; ***p<0.01

2.5 Activity #2 (Optional):

Time:

Discuss with a partner what the R^2 of each regression was, and what the value of the R^2 means for how we were able to explain the variation in exam scores.

∼ Questions? ∼

2.6 Summary Activity (5 Q's of Reflection):

1. What is the typical interpretation of the intercept?
2. Describe in your own words, one-three sentences, why we might not always be able to interpret the intercept.
3. Since $\bar{\hat{u}} = 0$, knowing the definition of $\hat{u}_i$, show that $\bar{Y} = \bar{\hat{Y}}$.
4. Give an example of when an X variable can be added to a regression to increase R^2 but might not actually be the variable driving variation in Y.
5. What questions do you still have regarding interpretation of regression output; specifically, questions on the coefficients or R^2 of the output?

3 Be Able To

- Interpret the coefficient of a regression output with a constant and a single X variable.
- Interpret the coefficient of regression output in a linear transformation of the X variable.
- Explain why the intercept might not be interpretable, and what to do if it is not.
- Prove that the mean of the residuals is zero.
- Prove that the sample correlation between X and residuals, and also $\hat{Y}$ and residuals therefore, is zero.
- Decompose TSS into RSS and ESS, and define R^2.
- Explain in words what R^2 is.
- Explain in words why R^2 doesn't necessarily matter.

4 Key Words

- Total Sum of Squares
- Residual Sum of Squares
- Explained Sum of Squares