Handout: Reading Regression Tables

When we look at a regression we are generally interested in two questions: What is the relationship between two things? How precise is the estimate of the relationship? We find the answer to the first question in the estimated "coefficient" and the answer to the second question is found in the "T-statistic."

First, let's focus on the estimated relationship. Before getting into an example, here are a few points to keep in mind.

- A positive regression coefficient is similar to two things being positively correlated (i.e. they move together).

- Conversely, a negative regression coefficient is similar to two things being negatively correlated (i.e. they move in opposite directions, when one goes up the other goes down).

Example: Say you were given the following table of results (these numbers are all made up) about the weight of tigers:

Dependent Variable: Weight of Tigers

Variable	Coefficient	Standard Error	T-Statistic
Constant	100	50	2
Age	10	2.4	0.5
Calories consumed (per day)	0.1	0.02	5
Length	5	2	2.5
Exercise hours (per day)	-20	15	1.33
Number of Whiskers	0.2	0.05	4

You can always rewrite a regression table as an equation. For example, based on the table above you could write:

Weight = 100 + 10 Age + 0.1 Calories + 5 Height – 20 Exercise + 0.2 Whiskers + error

Now say we have two tigers, Tigger and Tony.

	Tigger:	**Tony:**
Age	5	5
Calories consumed:	2000	1000
Length	5 feet	5 feet
Exercise Hours	2 hours	2 hours
Whiskers	5 whiskers	5 whiskers

Given our estimates from the table, what would we expect each of their weights to be?

Weight Tigger = 100 + 10*5 + 0.1*2000+ 5*5 – 20*2 + 0.2*5 = 336 pounds
Weight Tony = 100 + 10*5 + 0.1*1000+ 5*5 – 20*2 + 0.2*5= 236 pounds

What is the expected difference in Tigger and Tony's weights? You could estimate each tiger's weight using the formulas above and then take the difference: 336-236 = 100 pounds. However this isn't always convenient and there is a faster way to estimate this. Notice, the only difference between the two equations is the number of calories eaten per day. So you can just take the coefficient of calories (0.1) and multiply it by the difference in calories to get the expected difference in weights (see below).

Weight Tigger = 100 + 10*5 + 0.1*2000+ 5*5 – 20*2 +0.2*5 = 336 pounds
Weight Tony = 100 + 10*5 + 0.1*1000+ 5*5 – 20*2 + 0.2*5 = 236 pounds

---

Tigger – Tony = 0.1 *(2000-1000) = 100 pounds

Precision and Statistical Significance

While we've talked a lot about the estimates, we haven't looked at how precise they are. How big an error is "too big" depends on what you're estimating. If I tell you that my standard error is 5 you have no idea if I have a precise estimate. It depends on the size of what I'm estimating. For example, if I'm estimating United States' GDP it's an incredibly precise estimate (5 is tiny compared to $10 trillion, the size of GDP). However, if I'm talking about height, 5 feet is huge. If that is the error of your estimate, you have absolutely no idea how tall someone is. T-statistics provide a standardized way for researchers to communicate how precise their estimates are.

$$\text{T-statistic} = \frac{\text{Coefficient}}{\text{Standard Error}}$$

How big is "big enough?"

It depends on what you want to know and how precise you need to be. As a general rule of thumb – and in Econ 113 – we will say if the T-statistic is greater than 2, then the estimate is statistically significant (note if the t-statistic were less than negative 2 it would also be significant). We use 2 as a rule of thumb because you need a t-statistic at least that big to say that 95% of the time the coefficient is **not** just as likely to be 0 as it is to be number we estimated (remember a coefficient of 0 means that there is no relationship between the two variables). In other words, if the t-statistic is greater than or equal to 2 we can be fairly confident that our estimate tells us the sign (positive or negative) of the coefficient.

Back to our example, which coefficient is **not** statistically significant?

Answer: The coefficient on age has a t-statistic of 0.5, therefore it is not statistically significant.

Statistically Significant vs. Economically Significant

Statistically significance **<u>only</u>** tells you how precise your estimate is. It does not tell you anything about whether your estimate is important!!! We say something is statistically significant if it is precisely estimated, but we say it is economically significant if it is important.

- For example, in the table above we were able to very precisely estimate the relationship between whiskers and weight (t-statistic = 4). However, our estimates also tell us that the number of whiskers do not translate into much weight. 5 whiskers are only associated with 1 pound of weight, less than half of a percent of the weight of our smallest tiger, Tony. That's hardly important. Therefore, we would say that the relationship between tiger weight and whiskers is statistically significant, but it is not economically meaningful.

- Something can be economically significant even if it has a very small coefficient. For example, some theoretical models predict the sign of a coefficient. If we estimate the relationship and find the opposite sign (even if the coefficient is very small), that would be economically significant because it would call a theory into question.

- Similarly something can be statistically insignificant and economically significant. Just because something is important, does not mean that it is easy to estimate.